LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

numero 161
gennaio 1982
anno XV
volume XXVIII

L'intensificatore di immagine a microcanali

Un wafer di vetro perforato da milioni di tubi fotomoltiplicatori di elettroni è in grado di trasformare una figura confusa di radiazione elettromagnetica in un'immagine luminosa e nitida di tipo divisionista

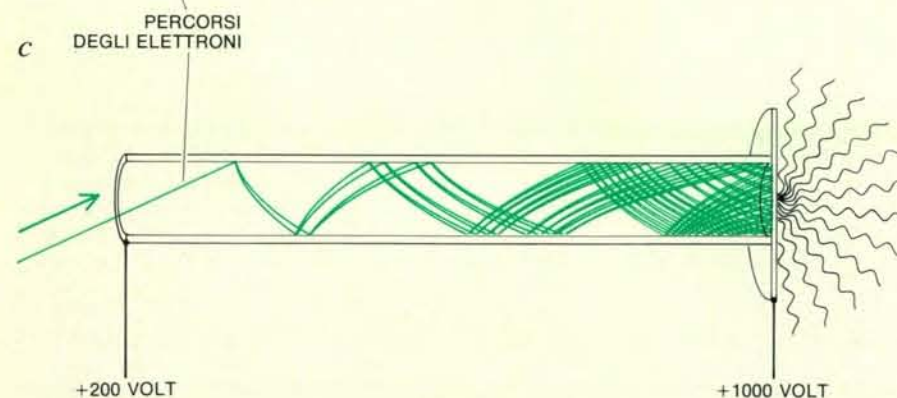
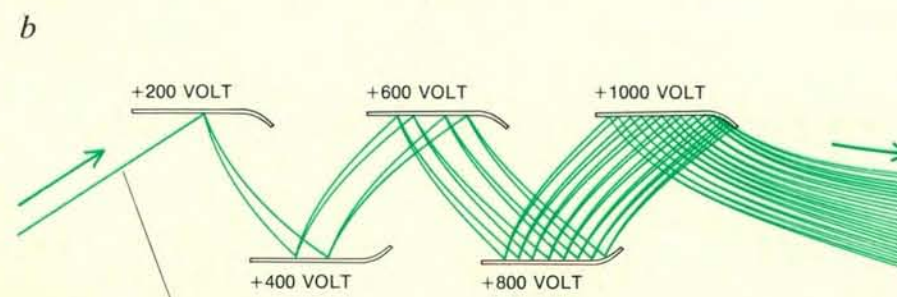
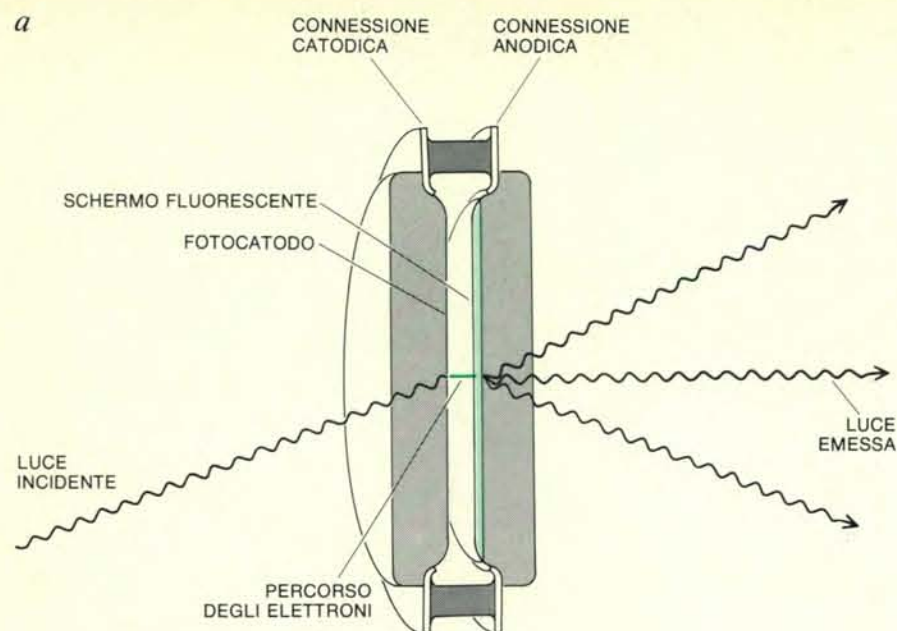
di Michael Lampton

Un amplificatore è un dispositivo che aumenta la potenza di un segnale, alterandone il meno possibile le altre caratteristiche. In generale si considera l'amplificazione come un processo applicabile solo a segnali che variano nel tempo, come per esempio l'uscita elettrica di un microfono. Un'immagine visiva costituisce un segnale che varia non solo nel tempo, ma anche nello spazio; per amplificare un segnale del genere è stato spesso necessario convertirlo in un'informazione variabile solo in funzione del tempo. Ciò si può ottenere esplorando l'immagine secondo un criterio sistematico, come avviene in una telecamera. Negli ultimi 25 anni i notevoli progressi compiuti dalla tecnologia del vetro hanno reso possibile raggruppare milioni di microscopici amplificatori disposti in schiere geometricamente ordinate. Ogni amplificatore della schiera è costituito da un tubo o canale di circa 15 micrometri di diametro in grado di esaltare la luminosità di una piccola, ma ben definita porzione dell'immagine. Quindi l'intera schiera di amplificatori, funzionanti contemporaneamente e in parallelo, si comporta come un intensificatore di luce e aumenta la luminosità di immagini deboli senza distruggere l'informazione spaziale contenuta nel segnale di ingresso. Una simile schiera di amplificatori microscopici è detta piastra a microcanali o piastra fotomoltiplicatrice.

Gran parte delle ragioni economiche che hanno condotto allo sviluppo della piastra a microcanali sono di origine militare. Un compatto intensificatore di immagini, comprendente una piastra fotomoltiplicatrice, è in grado di aumentare la luminosità di una scena fino a 10 000 volte, rendendo possibile la ricognizione notturna in condizioni di visibilità molto più ridotta di quanto non si riuscisse in passato. Occhiali, binocoli e strumenti analoghi dotati di intensificatori a microcanali sono diventati indispensabili per le operazioni militari notturne, per esempio per la guida di veicoli a fari spenti. La responsabilità principale per continuare lo sviluppo di questi dispositivi è stata affidata all'US Army Night Vision Laboratory di Fort Belvoir in Virginia. Al momento attuale il fabbisogno di intensificatori a microcanali richiede una produzione di circa 10 000 unità all'anno.

L'intensificatore a microcanali ha però applicazioni anche non militari. In astronomia può aumentare la luminosità di un'immagine fino a consentire la rivelazione di fotoni singoli. Ciò consente di registrare nel corso di alcuni minuti immagini che avrebbero altrimenti richiesto tempi di esposizione fotografica di durata proibitiva. Sono disponibili anche intensificatori di immagini per pazienti affetti da una forma di cecità notturna derivante da retinite pigmentosa, pur se il costo (intorno ai 2700 dollari negli Stati Uniti, pari a oltre tre milioni di lire) rappresenta una barriera per una loro più vasta diffusione.

Gli impieghi della piastra a microcanali non sono limitati all'intensificazione di immagini visibili: accoppiato con un adatto sistema elettro-ottico, il dispositivo può convertire un segnale bidimensionale nell'infrarosso vicino o nella regione dell'ultravioletto dello spettro elettromagnetico in un'immagine visibile. Può inoltre rivelare direttamente raggi X, ioni ed elettroni rendendo possibili numerosissime nuove applicazioni in vari campi: la progettazione di oscilloscopi, la microscopia elettronica e lo studio della composizione chimica dei materiali con metodi fotometrici. Una piastra a microcanali è impiegata nel telescopio a raggi X a bordo del satellite Einstein Observatory e negli spettrometri per l'ultravioletto installati a bordo dei Voyager in missione verso Giove e Saturno. L'European X-ray Observatory Satellite *(EXOSAT)*, che verrà lanciato nel 1982 dall'ESA, l'Agenzia spaziale europea, avrà a bordo alcuni rivelatori di raggi X dotati di piastre a microcanali.

Su quali principi si basa il funzionamento della piastra a microcanali? È opportuno iniziare la spiegazione descrivendo un intensificatore di immagine non basato su tale concetto. Si tratta di un dispositivo che sfrutta l'interazione del-

Gli intensificatori di immagine sfruttano l'effetto fotoelettrico, per mezzo del quale l'energia dei fotoni incidenti, o quanti di radiazione elettromagnetica, viene convertita nell'energia degli elettroni in moto. Gli elettroni sono particelle dotate di carica elettrica, al contrario dei fotoni, e quindi la loro energia può essere accresciuta mediante accelerazione da parte di un campo elettrico. Quando i fotoni colpiscono una piastra metallica fotosensibile posta nel vuoto provocano l'emissione di elettroni; questi sono accelerati verso una piastra fluorescente carica positivamente e qui la loro energia è nuovamente convertita sotto forma di luce visibile. In un tubo intensificatore di immagine a due elettrodi *(a)* per ogni fotone incidente viene emesso almeno un elettrone secondario; l'elettrone, accelerato dal campo, deve dirigersi direttamente nello spazio fra i due elettrodi, se non si vuole che l'immagine diventi confusa. Per conservare la nitidezza dell'immagine, gli elettrodi devono essere perciò molto vicini fra loro. Il guadagno ottenibile è limitato dalla tensione che può essere applicata fra gli elettrodi senza che si inneschi l'autoscarica. Gli elettroni che colpiscono certi materiali possono provocare l'emissione di ulteriori elettroni: si può indurre così una cascata di elettroni applicando tensioni crescenti a diversi elettrodi posti in cascata *(b)*. Sebbene un moltiplicatore di questo genere possa dare un elevato guadagno, la dispersione degli elettroni in cascata distrugge l'informazione relativa alla distribuzione spaziale dei fotoni incidenti. Una via per ottenere nel contempo alto guadagno e buona risoluzione spaziale è quella di avere un moltiplicatore di elettroni per ogni pixel, ossia per ogni elemento di quadro dell'immagine, ma, in questo caso, i moltiplicatori devono essere di dimensioni piccolissime. Questo può essere realizzato con il moltiplicatore di elettroni a canali continui *(c)*, che funziona in modo molto simile all'elettrodo moltiplicatore, ma confina la cascata di elettroni al foro del canale.

l'energia elettromagnetica con la materia, fenomeno spiegato per la prima volta da Einstein nel 1905. L'interazione in questione è l'effetto fotoelettrico: questo è stato uno dei primi processi fisici che ha fatto sorgere l'idea che l'energia varia in quanti, ossia non con continuità, ma in «pacchetti» discreti.

Quando un fotocatodo, o elettrodo negativo, viene esposto a radiazione elettromagnetica di lunghezza d'onda inferiore a un certo valore critico, emette una corrente di elettroni che possono essere raccolti su un anodo, o elettrodo positivo. Per molti metalli la lunghezza d'onda critica si trova nella regione dell'ultravioletto, mentre per alcune sostanze è nella regione visibile dello spettro o in quella dell'infrarosso vicino. La quantità di carica trasferita è proporzionale all'intensità della radiazione, ma la possibilità del trasferimento o meno della carica dipende solo dalla lunghezza d'onda della radiazione incidente.

Einstein spiegò il fenomeno supponendo che la liberazione di un elettrone facente parte di un atomo di materia richiede una certa quantità minima ben definita di energia. Inoltre immaginò che l'energia occorrente per liberare l'elettrone, se fornita da radiazione elettromagnetica, può essere assorbita solo in quantità discrete sotto forma di quanti detti fotoni. Dai precedenti studi di Max Planck si sapeva che l'energia di un fotone è inversamente proporzionale alla lunghezza d'onda della radiazione; pertanto un fotone di lunghezza d'onda maggiore della lunghezza d'onda critica possiede un'energia insufficiente per tirar fuori un elettrone dal metallo.

Un intensificatore di immagine sfrutta l'effetto fotoelettrico per convertire il flusso di fotoni, contenuto nella radiazione elettromagnetica, in un flusso di elettroni. Ogni fotone può stimolare l'emissione di almeno un elettrone e potrebbe quindi sembrare che tale procedimento non consenta alcun guadagno; ma ciò che rende possibile l'amplificazione dell'immagine è la carica elettrica dell'elettrone. Quando gli elettroni sono sottoposti all'azione di un campo elettrico possono essere accelerati, e in questo modo aumenta l'energia cinetica che possiedono. In tale maniera l'energia disponibile per la formazione di un'immagine può venire moltiplicata per fattori sensibilmente maggiori di uno.

Nella forma più semplice un intensificatore di immagine è costituito da un tubo a vuoto con due elettrodi connessi a una sorgente esterna di alimentazione, ma non contenente alcuna piastra a microcanali. L'immagine viene focalizzata sul fotocatodo: ogni areola dell'elettrodo emette elettroni in un numero che è determinato dalla luminosità dell'immagine in quel punto. Ogni elettrone è poi accelerato da un campo elettrico generato da un potenziale di polarizzazione di diverse migliaia di volt applicato fra i due elettrodi. Gli elettroni accelerati si spostano verso l'anodo seguendo traiettorie approssimativamente rettilinee. Il flusso di elettroni che raggiunge l'anodo corrisponde quindi alle variazioni di luminosità dell'immagine sul fotocatodo. Il potere risolutivo spaziale dell'immagine è ben mantenuto poiché anodo e catodo sono disposti molto vicini fra loro.

L'uscita dell'intensificatore di immagine è visibile poiché l'anodo è costituito da materiale fluorescente. Quando un elettrone colpisce l'anodo, cede la propria energia e quindi eccita e ionizza atomi nel materiale di cui è costituito l'anodo stesso. Una frazione di quest'energia viene restituita sotto forma di fotoni di luce visibile, in quanto gli elettroni appartenenti agli atomi eccitati tendono a ritornare ai loro livelli energetici originali. I fotoni emessi secondo il meccanismo descritto costituiscono l'immagine visibile sull'anodo fluorescente.

Ho sottolineato prima come ogni fotone che colpisce il catodo possa liberare almeno un elettrone; in effetti il rendimento del fotocatodo è di appena il 10 per cento e ciò significa che solo un fotone su dieci stimola l'emissione di un elettrone. Inoltre solo il 30 per cento dell'energia convogliata sull'anodo dagli elettroni accelerati viene convertita in luce. Malgrado le perdite così rilevanti il tubo presenta un guadagno netto (cioè una amplificazione della luminosità dell'immagine) perché l'energia di ogni elettro-

Fotografia di una scena notturna con illuminazione stellare eseguita con macchina da 35 millimetri dotata di intensificatore d'immagine a microcanali. La luminosità della fotografia è molto simile a quella della scena come apparirebbe vista dall'occhio umano attraverso l'intensificatore. Il colore verde del cielo e delle cime degli alberi è il colore della luce fluorescente emessa dallo schermo su cui è visibile l'immagine amplificata e corrisponde alla lunghezza d'onda della luce alla quale l'occhio è più sensibile, cioè circa 540 nanometri. La fotografia è stata ripresa con una pellicola a colori per luce diurna avente sensibilità ASA 400, con un tempo di esposizione di un quarto di secondo e apertura del diaframma *f*/2,8. Senza intensificatore di immagine una fotografia del genere richiederebbe un tempo di esposizione compreso fra i quattro e gli otto minuti. L'intensificatore è stato fornito da Paul Lighty della Electro-Optical Products Division della ITT di Roanoke in Virginia.

ne emesso può essere aumentata di alcune migliaia di volte.

Si potrebbe ritenere che un segnale di immagine possa essere amplificato fino a raggiungere un qualsivoglia livello, semplicemente aumentando il potenziale di polarizzazione che accelera gli elettroni. Purtroppo un intensificatore di immagine a due elettrodi non può funzionare quando la tensione applicata supera un certo valore massimo, raggiunto il quale si comincia a verificare un flusso spontaneo di cariche nel breve spazio fra fotocatodo e anodo; tale flusso maschera del tutto la ben più debole corrente di elettroni che trasmette l'immagine. L'effetto potrebbe essere ridimensionato progettando tubi con una maggiore distanza fra catodo e anodo, ma in tal caso il potere risolutivo spaziale dell'immagine risulterebbe deteriorato a causa di piccole e casuali deviazioni degli elettroni dal percorso rettilineo. Se gli elettroni vengono focalizzati con campi elettrici e magnetici, gli elettrodi possono essere disposti a maggiore distanza e sottoposti a tensione più elevata aumentando di conseguenza il guadagno di luminosità dell'immagine. I campi elettrici e magnetici devono tuttavia essere controllati con cura, ma anche in tal caso questi dispositivi di focalizzazione possono introdurre nell'immagine distorsioni geometriche o disuniformità di luminosità. Inoltre, per molte applicazioni nei campi della fisica e dell'astronomia persino gli intensificatori dotati dei più complessi dispositivi di focalizzazione degli elettroni offrono un guadagno insufficiente.

Il fatto che ogni fotone incidente possa determinare al massimo il trasferimento all'anodo di un elettrone soltanto è una limitazione fondamentale di questi tubi intensificatori di immagine. In certi materiali però gli elettroni possono essere liberati anche da un elettrone incidente, in modo analogo a quanto viene fatto da un fotone incidente. L'effetto è chiamato «emissione secondaria superficiale» e può venire sfruttato in un intensificatore di immagine per moltiplicare il numero degli elettroni che vanno a raggiungere l'anodo.

I moltiplicatori di elettroni sono stati impiegati fin dall'inizio degli anni cinquanta con la funzione di componenti amplificatori nei tubi fotomoltiplicatori. A elettrodi metallici denominati dinodi, o catodi secondari, vengono applicati potenziali positivi via via crescenti; i dinodi sono disposti in modo che gli elettroni vengano diretti da un elettrodo al successivo. Un elettrone incidente che colpisce il primo dinodo provoca l'emissione dal metallo di alcuni elettroni dotati di bassa energia. Gli elettroni secondari vengono accelerati verso il secondo dinodo, dove si moltiplicano di nuovo. Successive moltiplicazioni determinano una crescita esponenziale della carica elettrica liberata dall'elettrone iniziale. Poiché tutti gli elettroni sono accelerati dal campo, l'energia combinata degli elettroni che alla fine sono raccolti dall'anodo può risultare enormemente superiore all'energia che è posseduta da un singolo elettrone accelerato.

Un moltiplicatore a dinodi può fornire un'amplificazione molto elevata, ma solo a spese del potere risolutivo. La geometria di un dinodo non assicura che gli elettroni si spostino secondo traiettorie rettilinee: ne consegue perciò che gli elettroni emessi da una data zona del fotocatodo non raggiungono di necessità la corrispondente zona dell'anodo. La piastra a microcanali invece è un dispositivo che offre nello stesso tempo il guadagno di un moltiplicatore di elettroni e il potere risolutivo di un intensificatore di immagine. Essa è costituita da milioni di moltiplicatori, ciascuno indipendente dagli altri, che sono assiemati in una schiera bidimensionale. Non è importante ottenere un certo grado di risoluzione spaziale all'interno di ogni singolo moltiplicatore; ogni moltiplicatore corrisponde piuttosto a un singolo pixel, ossia l'elemento unitario dell'immagine, e l'immagine intensificata è formata da un insieme di piccoli punti di luminosità variabile.

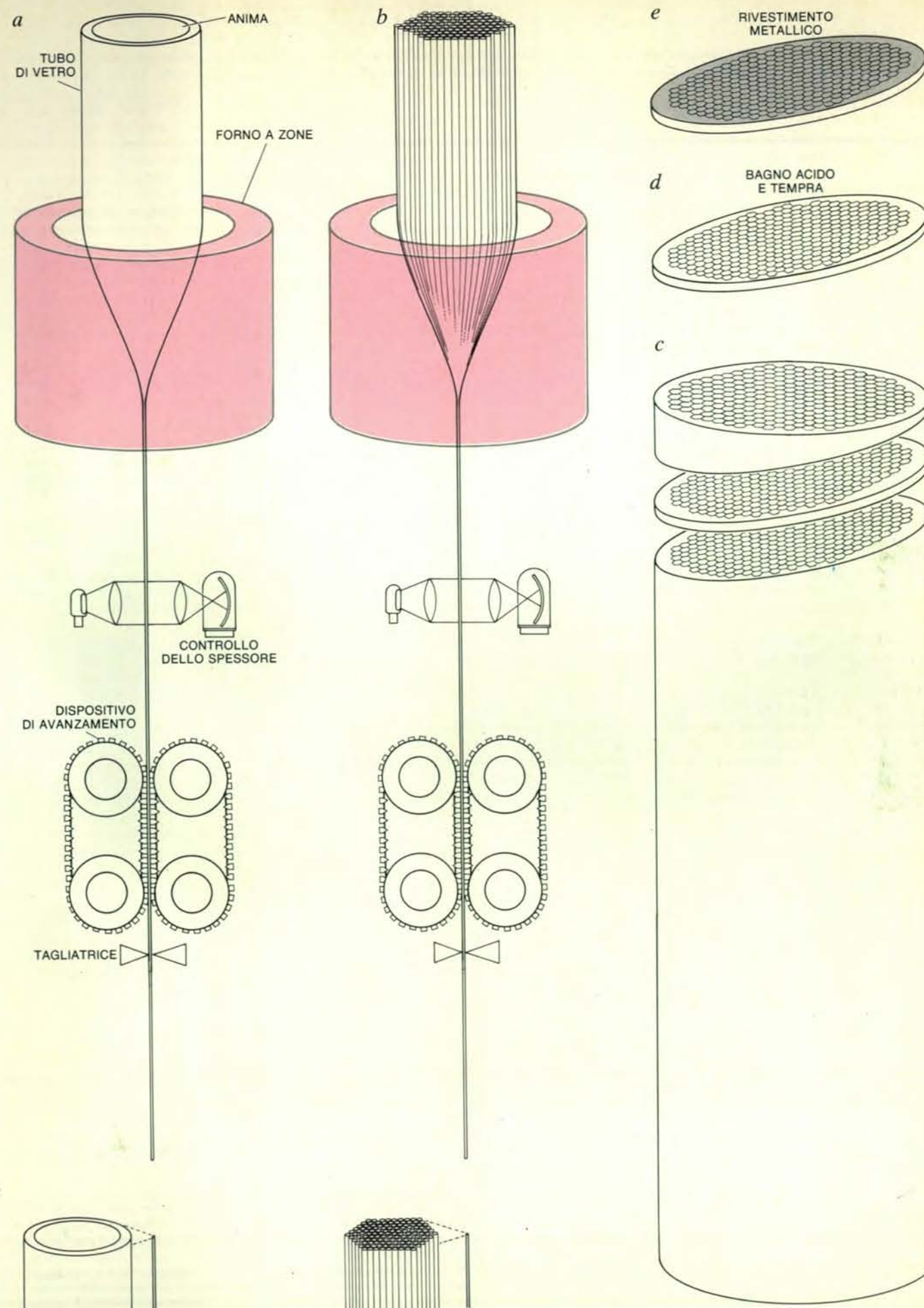
La fase decisiva nello sviluppo della piastra a microcanali è stata l'invenzione del moltiplicatore di elettroni a canale continuo. Si tratta di un tubo in vetro collegato a ciascuna estremità a una sorgente di energia elettrica, alimentato da un potenziale di polarizzazione di circa 1000 volt e collocato nel vuoto. Il dispositivo funziona, però, in modo molto simile a un moltiplicatore a dinodi. Un elettrone incidente può avviare una cascata di elettroni secondari che cresce esponenzialmente man mano che avanza lungo il tubo. La cascata è del tutto confinata all'interno del tubo o canale. Moltiplicatori a canale continuo di questo tipo sono stati studiati per la prima volta presso i Bendix Research Laboratories di Southfield, Michigan, da George W. Goodrich, James R. Ignatowski e William C. Wiley fra il 1959 e il 1961.

I materiali di cui sono costituiti i canali moltiplicatori devono soddisfare due requisiti. In primo luogo la parete del canale deve essere in grado di emettere più elettroni di quanti ne assorba. Per i livelli pratici di energia degli elettroni esistono diversi materiali, fra cui molti vetri, che emettono in media due elettroni per ogni elettrone incidente. In secondo luogo la conducibilità elettrica del materiale deve essere prevedibile e controllabile, in modo che la carica allontanata dalla parete del canale possa essere sostituita ristabilendo un campo elettrico uniforme. Il soddisfacimento del secondo requisito ha rappresentato una notevole difficoltà per i primi sperimentatori.

In un vetro la conduzione di una corrente elettrica può essere assicurata dalla diffusione sia di elettroni liberi sia di ioni (cioè di atomi carichi). Per ottenere una conduzione controllata è necessario ridurre al minimo la conducibilità ionica. La diffusione degli ioni è un tipo di elettrolisi che dà luogo a un'indesiderata attività chimica in corrispondenza degli elettrodi, come l'esaurimento dei metalli alcalini e la polarizzazione di cariche elettriche all'interno del vetro. Occorre perciò assicurare la conduzione dell'elettricità esclusivamente mediante il moto di elettroni liberi.

Nella realizzazione di canali moltiplicatori sono stati impiegati molti vetri di differenti composizioni, ma il vetro di gran lunga più adatto si è dimostrato quello costituito da una miscela composta da ossido di piombo per il 50 per cento circa, da biossido di silicio per il 40 per cento e da minori quantità di diversi ossidi alcalini per il rimanente. In assenza di ulteriori trattamenti, un vetro di questo tipo presenta un'elevata resistività elettrica, ovvero in altre parole una bassa conducibilità. Gli atomi di piombo presenti nel vetro sono stabili sotto forma sia di atomi neutri sia di ioni positivi che risultano dopo la cessione di elettroni nella formazione del legame ionico di valenza con l'ossigeno. La conducibilità può essere migliorata eliminando l'ossigeno legato con il piombo per mezzo di una reazione chimica di riduzione.

La riduzione può essere eseguita sulla superficie del vetro per riscaldamento in atmosfera di idrogeno. Dopo diverse ore a 400 gradi centigradi l'ossido di piombo risulta ridotto per una profondità di vari decimi di micrometri mentre la maggior parte del piombo ridotto ha abbandonato la superficie per evaporazione. Sulla superficie il piombo residuo si raccoglie per coalescenza in aggregati metallici, che impartiscono al vetro un caratteristico

Il processo a doppia trafilatura dà origine a piastre a microcanali con fibre ottenute per «trafilatura» da una massa o *boule* di vetro caldo. Le fibre conservano la geometria della sezione trasversale della massa di partenza, anche se su scala ridotta. La fibra viene forata e nel tubo cilindrico di vetro risultante, del diametro di 50 millimetri circa, viene inserita un'«anima» di vetro avente composizione diversa, che può essere eliminata, alla fine del processo con acido. Il tubo è riscaldato, trafilato fino a raggiungere all'incirca lo spessore di un millimetro *(a)* e poi tagliato in spezzoni lunghi 15 centimetri. Diverse migliaia di spezzoni sono raggruppati in un fascio esagonale avente uno spessore di circa 50 millimetri. Il blocco così formato è poi riscaldato di nuovo e trafilato fino a ridurlo a uno spessore grosso modo di un millimetro *(b)*. La fibra che si ottiene è una copia geometricamente simile all'esagono da 50 millimetri e con la struttura di sottili tubi in vetro raggruppati. Le fibre esagonali sono poi fuse insieme e tagliate ad angolo obliquo per ottenere dei wafer spessi circa un millimetro *(c)*. I wafer sono successivamente immersi in un bagno d'acido per eliminare l'anima di vetro e riscaldati in atmosfera di idrogeno per impartire alle pareti interne dei microcanali la voluta conducibilità elettrica *(d)*. Su entrambe le facce del wafer viene fatto depositare, per evaporazione, del metallo su cui possono essere effettuate le necessarie connessioni elettriche *(e)*. Le dimensioni dei tubi raggruppati nell'esagono e degli esagoni raggruppati sul wafer finito sono state molto aumentate, nel disegno, per chiarezza di rappresentazione.

colore nero. Gli elettroni, che in origine formavano il legame ionico dell'ossido di piombo, rimangono liberi e possono condurre una corrente elettrica: lo strato superficiale del vetro è così diventato un semiconduttore con resistività compresa fra 10^8 e 10^{14} ohm per quadrato, una gamma di valori adatta per la costruzione di moltiplicatori a canale e di piastre a microcanali. (La resistività superficiale si esprime in ohm per quadrato, senza specificare le dimensioni del quadrato, in quanto la resistività di una regione quadrata sulla superficie di una sostanza uniforme non dipende dalle dimensioni del quadrato stesso.)

Le prime schiere sperimentali di moltiplicatori a canale continuo vennero montate legando assieme alcune dozzine di canali moltiplicatori realizzati con un vetro a basso punto di fusione. Ogni canale aveva un diametro di circa un millimetro e quindi il potere risolutivo di questo strumento primitivo era piuttosto grossolano. Inoltre, per riprodurre un'immagine di buona qualità è necessario disporre di un numero ben maggiore di elementi amplificatori. Presso i Bendix Research Laboratories fu laboriosamente montata a mano una schiera esagonale di circa 5000 canali i cui assi distavano fra loro di 0,15 millimetri. Sebbene la risoluzione per punti di una schiera del genere sia superiore a quella in generale raggiunta nella riproduzione di fotografie su quotidiani e periodici, l'area complessiva dell'immagine è di appena un centimetro quadrato circa.

Potrebbe sembrare che il risultato degli esperimenti condotti con dispositivi aventi le dimensioni dei primi strumenti non possa essere estrapolata a schiere microscopiche di canali moltiplicatori in grado di trasmettere immagini di alta qualità. L'estrapolazione è invece possibile perché per una data differenza di potenziale applicata, il guadagno del canale moltiplicatore non dipende dalla dimensione di quest'ultimo: aumentando infatti la lunghezza del canale ogni elettrone è soggetto più a lungo all'influenza della forza esercitata dal campo elettrico e ha così un tempo maggiore per acquisire energia. Se al contrario si riduce la lunghezza del canale il campo elettrico è più intenso e la forza che agisce sull'elettrone varia in modo inversamente proporzionale alla distanza fra gli elettrodi. I due effetti si elidono a vicenda con precisione.

Le schiere fatte a mano permisero di stabilire che i singoli canali danno ciascuno una risposta indipendente dalle altre. Si trovò tuttavia che la qualità dell'immagine è molto sensibile alle variazioni di diametro dei canali e in qualche modo

a

FILO METALLICO SOTTILE

RIVESTIMENTO DI VETRO

AVVOLGIMENTO SU BOBINA

TAGLIO IN SEZIONI

TAGLIO IN WAFER

b

BAGNO ACIDO

FUSIONE

TAGLIO IN WAFER

Per la fabbricazione di sottili tubi di vetro e per la loro sistemazione in grandi schiere sono state suggerite molte tecniche, rivelatesi non soddisfacenti prima che si riuscisse a sviluppare un metodo pratico. Secondo una delle tecniche *(a)*, un filo veniva ricoperto di vetro fuso; dopo il raffreddamento la fibra ottenuta era raccolta su una bobina e il rocchetto di fibre densamente avvolte che ne risultava era tagliato in sezioni, da cui venivano ricavati dei sottili wafer. Infine si eliminava l'anima, costituita dal filo, per immersione in un bagno acido, che lasciava il wafer formato da piccoli tubi di vetro. Risultò difficile conseguire la necessaria uniformità delle fibre durante l'avvolgimento; e inoltre il processo non è conveniente dal punto di vista economico per la necessità di manipolare fibre microscopiche fin dalla prima fase della lavorazione. Una seconda tecnica *(b)* impiegava la fotolitografia per incidere sottili lastre di vetro, che venivano poi raggruppate, fuse e tagliate in wafer. Nel corso del processo di incisione non era però possibile controllare la profondità e la larghezza dei canali entro le tolleranze volute.

anche alle disuniformità nella spaziatura fra i canali. Agli inizi degli anni sessanta la principale difficoltà per lo sviluppo delle piastre a microcanali fu l'ideazione di tecniche di fabbricazione che consentissero di mantenere un'elevata uniformità nel diametro dei canali.

Una tecnica chiamata «processo ad anima metallica», fu inventata presso i Mullard Research Laboratories di Londra e successivamente sviluppata dal Laboratorio di elettronica e di fisica applicata nei pressi di Parigi. Un sottile e uniforme filo metallico viene dapprima rivestito di vetro per canale moltiplicatore, reso plastico mediante riscaldamento e successivamente avvolto su un tamburo poligonale. Quando l'avvolgimento ha raggiunto un certo spessore viene tagliato in blocchi, che sono poi affiancati uno all'altro e fusi in una massa o *boule* di vetro moltiplicatore nella quale risultano imprigionate molte migliaia di fili. La massa viene poi tagliata in fette sottili e attaccata con acido diluito per eliminare i fili metallici. I wafer di vetro risultanti sono attraversati da migliaia di canali di diametro estremamente uniforme.

Il processo ad anima metallica ha due grossi svantaggi. Anzitutto, pur essendo il diametro dei canali costante, non dà luogo a una spaziatura uniforme: è infatti difficile avvolgere il fragile filo rivestito di vetro su un tamburo con uniformità perfetta. Secondariamente la fibra di vetro ha, durante tutte le fasi della fabbricazione, lo stesso microscopico diametro finale e quindi le macchine rivestitrici e avvolgitrici, che hanno un costo molto elevato, possono lavorare solo un piccolo volume di materiale nell'unità di tempo, il che si è dimostrato un serio ostacolo, sul piano economico.

Un'altra tecnica, denominata «processo a piastra scanalata», è stata studiata negli Stati Uniti e in Svizzera. Minuti solchi paralleli vengono incisi in gran numero mediante un processo fotolitografico sulle superfici di molte sottili piastre di vetro che sono poi affiancate e fuse per formare il blocco da cui si tagliano i wafer a microcanali. Il metodo sembrava a prima vista promettente poiché la spaziatura fra i solchi può essere controllata con elevata precisione durante l'incisione fotolitografica. Si possono poi realizzare anche canali curvi o a zig-zag, riuscendo così a superare alcune limitazioni di guadagno intrinseche alla geometria dei canali rettilinei. Il controllo dello spessore e della profondità dei solchi durante i processi di incisione e di fusione si è dimostrato piuttosto difficile e le conseguenti disuniformità di guadagno hanno avuto come conseguenza finale l'abbandono del processo.

La tecnica che si è dimostrata più adatta dal punto di vista economico per la fabbricazione in massa delle piastre a microcanali è detta «processo a doppia trafilatura». Esso è derivato dai metodi di trafilatura o di tiraggio del vetro in fibre di sezione microscopica; tali metodi cominciarono a essere sviluppati dagli artigiani dell'antico Egitto durante la Diciottesima Dinastia (1570-1320 a.C.). Certi vetri di composizione speciale e che sono lavorabili in una vasta gamma di temperature hanno la proprietà di conservare la loro sezione trasversale durante il riscaldamento e la trafilatura. I ricercatori dei Bendix sono stati i primi a sfruttare questa proprietà per la fabbricazione di piastre a microcanali nel 1960 e da allora il metodo è stato adottato anche da altri fabbricanti.

La piastra a microcanali finita è un sottile wafer che può essere montato in varie versioni di intensificatori di immagine. La piastra a microcanali standard da 25 millimetri qui riprodotta è costituita da qualcosa come tre milioni di microcanali capaci di moltiplicare la luminosità di un'immagine per un fattore di 10 000 circa. La piastra a microcanali che compare nella fotografia è stata fabbricata dalla Galileo Electro-Optics Corporation di Sturbridge, Massachusetts.

Nel processo a doppia trafilatura il vetro del canale moltiplicatore viene fuso o estruso in forma di lingotto cilindrico e poi raffreddato e lavorato fino a ottenere una barra uniforme del diametro di parecchi centimetri. La barra viene forata lungo l'asse e nel foro si inserisce di solito un'altra barra costituita da un vetro di composizione differente. Alla fine del processo l'«anima» costituita dalla seconda barra viene rimossa per immersione in un bagno di acido caldo diluito, ma durante le fasi intermedie della fabbricazione l'anima supporta il cilindro esterno in modo tale che i microcanali finiti risultino di diametro e forma pressoché uniformi.

Durante la prima trafilatura il cilindro di vetro è tenuto sospeso verticalmente in un forno a zone, ove la temperatura è controllabile punto per punto. Il fondo del cilindro è riscaldato fino a circa 500 gradi centigradi; una goccia di vetro fuso discende dal forno tenuta in sospensione da una fibra di vetro, cui una filiera impartisce un diametro di un millimetro circa. Quando la fibra è scesa fino a parecchi metri al di sotto del forno, si è raffreddata quanto basta per essere maneggiata da una macchina di avanzamento che regola la velocità della trafilatura. La regolazione della velocità assicura a sua volta un controllo di precisione del diametro della fibra. Al di sotto della macchina di avanzamento la fibra è tagliata in spezzoni lunghi circa 15 centimetri, spezzoni che sono infine riuniti insieme a migliaia in un fascio di forma esagonale.

La seconda trafilatura è simile alla prima. Il fascio esagonale di fibre è sospeso e riscaldato in un forno a zone e trafilato sotto forma di fibra composta esagonale avente dimensione trasversale di un millimetro circa. La spaziatura fra i singoli cilindri è così ridotta a pochi centesimi di millimetro, ossia la spaziatura finale nella piastra a microcanali. La fibra di vetro trafilata è nuovamente tagliata in segmenti che sono raggruppati e fusi sotto vuoto per formare una *boule* di solidi esagonali che può avere un diametro fino a

Per la trasmissione di immagini di qualità elevata l'uniformità delle dimensioni e della spaziatura dei microcanali è critica. Una piastra a microcanali ottenuta da fibre in vetro trafilato presenta una buona uniformità, come si può vedere nelle due microfotografie. In quella in alto, ripresa da Adolf R. Asam della Electro-Optical Products Division della ITT, la piastra a microcanali è ingrandita 110 volte e ne risulta chiaramente visibile la struttura secondaria esagonale. La microfotografia in basso, eseguita dall'autore, mostra una piastra a microcanali ingrandita 650 volte: in essa i microcanali, che hanno un diametro di 40 micrometri, sono spaziati di 50 micrometri.

125 millimetri e incorporare milioni di microcanali.

Le piastre a microcanali sono ottenute tagliando la *boule* in wafer spessi circa un millimetro, le cui facce vengono poi lucidate. Il taglio è di solito eseguito obliquamente rispetto all'asse dei microcanali in modo che nella piastra finita gli elettroni incidenti urtino la parete del canale vicino all'ingresso anziché viaggiare in linea retta verso l'anodo attraversando tutto il canale. Il wafer può essere molato, perforato e lavorato a caldo per soddisfare requisiti particolari. Nelle fasi finali il vetro dell'anima viene sciolto mediante bagno acido e la piastra viene riscaldata in atmosfera di idrogeno per ridurre l'ossido di piombo e impartire alla superficie del vetro il valore voluto di conducibilità elettrica. La finitura del wafer è realizzata mediante riporto per evaporazione su entrambe le facce di una pellicola metallica, per poter eseguire connessioni elettriche per tutti i canali.

Esistono due tipi fondamentali di intensificatori di immagine dotati di una piastra a microcanali. L'intensificatore a focalizzazione di prossimità assomiglia a un semplice intensificatore a due elettrodi nel quale sia stata inserita una piastra a microcanali. L'immagine incidente è messa a fuoco sul fotocatodo; gli elettroni emessi da quest'ultimo sono accelerati nel breve percorso fino alla piastra a microcanali e gli elettroni che ne emergono in numero accresciuto vengono ulteriormente accelerati nel percorso leggermente più lungo del precedente fino a quando raggiungono uno schermo fluorescente. Un tubo del genere è compatto, esente da distorsioni e insensibile ai campi magnetici. Inoltre, l'immagine in uscita è piana e sempre a fuoco, come avviene nel caso della stampa a contatto su carta fotografica.

L'intensificatore a focalizzazione di prossimità si presta in modo particolare per eseguire fotografie di fenomeni molto rapidi. La piastra a microcanali intensifica la debole immagine ottenuta durante il breve tempo di esposizione fotografica. Il dispositivo può anche servire come otturatore fotografico ultrarapido: la differenza di potenziale controllata dall'esterno fra il fotocatodo e la superficie frontale della piastra fornisce un mezzo rapido per abilitare e disabilitare l'intensificatore. Un intensificatore a focalizzazione di prossimità può rispondere a un impulso elettrico di abilitazione di appena alcuni nanosecondi: esso consente quindi di eseguire fotografie nitide di fenomeni aventi una durata estremamente breve.

Il secondo tipo di intensificatore di immagine impiega un campo elettrostatico non uniforme per focalizzare sulla piastra a microcanali gli elettroni emessi dal fotocatodo. Il tubo, come una lente ottica, capovolge l'immagine e può anche ingrandirla; in alcuni casi il dispositivo può funzionare perfino come uno zoom, ossia un obiettivo a focale variabile che regola contemporaneamente fuoco e ingrandimento. A causa del capovolgimento dell'immagine, l'intensificatore a focalizzazione elettrostatica è particolarmente utile se viene usato insieme a dispositivi ottici che capovolgono anch'essi l'immagine. La piastra a microcanali può sostituire un sistema ottico complesso, come quello dei binocoli che raddrizzano l'immagine capovolta che si forma attraverso un obiettivo normale.

Sebbene la piastra a microcanali sia stata applicata principalmente all'intensificazione di immagine, può amplificare qualsiasi forma di radiazione capace di dare origine a una cascata di elettroni nei canali. Ioni e elettroni, come pure fotoni di lunghezza d'onda nell'ultravioletto e nella regione X dello spettro, sono in grado di dare origine a cascate di elettroni nei microcanali. Questa forma di radiazione è pertanto percepibile senza l'ausilio di un fotocatodo. La radiazione è messa a fuoco sulla superficie frontale della piastra e gli elettroni emessi sono rivelati sulla superficie posteriore. Se la piastra è dotata di uno schermo anodico fluorescente, può funzionare come convertitore di immagine, rendendo percepibile all'occhio la distribuzione spaziale della radiazione che si trova al di fuori della parte visibile dello spettro.

Le piastre a microcanali, che funzionano sia come convertitori sia come intensificatori, sono state incorporate in diversi strumenti di laboratorio: oscilloscopi ad alta velocità, microscopi elettronici a trasmissione, microscopi ionici e spettrometri di massa. In queste applicazioni il guadagno della piastra a microcanali rende possibile la registrazione di eventi brevissimi, o che hanno una debole luminosità oppure uno scarso contrasto. Inoltre consente di far funzionare lo strumento con una corrente del fascio minore, ciò che costituisce talvolta un vantaggio importante. In un microscopio elettronico, per esempio, si possono ridurre i danni al materiale in esame provocati dal bombardamento elettronico.

Quattro limitazioni fisiche principali riducono le prestazioni delle usuali piastre a microcanali. La prima consiste nel fatto che il segnale medio di uscita non può superare la corrente massima tollerabile dalle pareti del microcanale. Quando il flusso di elettroni è troppo intenso, la carica elettrica asportata dal vetro non viene sostituita immediatamente e il campo si modifica con conseguente diminuzione del guadagno complessivo. Zone luminose di un'immagine possono risultare così saturate, con perdita di contrasto nell'area corrispondente. Questo effetto non è desiderabile in astronomia, ove è richiesta una gamma dinamica ben calibrata, ma può tornare utile per i dispositivi di visione notturna nei quali un forte contrasto nelle parti più luminose dell'immagine può diminuire la sensibilità visiva dell'osservatore nei confronti delle zone più scure.

Una seconda limitazione dei moltiplicatori è il fenomeno chiamato retroazione ionica. Quando un canale funziona con alto guadagno, gli atomi di gas presenti nel canale possono essere ionizzati da collisioni con gli elettroni in cascata. (Anche con il vuoto più spinto raggiungibile sono sempre presenti almeno alcuni atomi di gas.) Uno ione dotato di carica positiva formato in questo modo viene accelerato dal campo elettrico verso l'estremità d'ingresso del tubo ove può colpire la parete del canale e iniziare una nuova cascata. Questa a sua volta può provocare un'ulteriore ionizzazione e gli elettroni addizionali che ne derivano tendono a mascherare il segnale. Inoltre gli ioni che colpiscono il fotocatodo ne accorciano la vita, benché danni di questo genere si prevengano ora di norma mediante il deposito sulla superficie di ingresso della piastra di una sottile pellicola di ossido di alluminio, che risulta permeabile ai fotoelettroni, ma non agli ioni.

Un accorgimento per minimizzare la retroazione ionica è quello di realizzare canali curvi o a zig-zag. Gli elettroni possono così dirigersi facilmente verso l'anodo, mentre il movimento degli ioni

L'incorporamento delle piastre a microcanali nei tubi intensificatori di immagine viene effettuato secondo due configurazioni principali. In entrambe le configurazioni la piastra è serrata fra un fotocatodo, ossia un elettrodo a carica negativa, e un anodo fluorescente, cioè un elettrodo a carica positiva. Le tensioni degli elettrodi sulla piastra a microcanali hanno valori intermedi a quelle del fotocatodo e dell'anodo: nel diagramma i potenziali dei quattro elettrodi sono positivamente crescenti da sinistra a destra. L'intensificatore a focalizzazione di prossimità *(diagramma in alto)* è analogo a un intensificatore a due elettrodi. Gli elettroni emessi dal fotocatodo attraversano il breve percorso fino alla piastra a microcanali, dove sono accelerati e moltiplicati. Gli elettroni che emergono dalla piastra sono raccolti dall'anodo, dopo aver attraversato una breve distanza. La distribuzione spaziale dell'immagine di ingresso viene conservata perché i due percorsi fra gli elettrodi e la piastra sono tanto brevi da minimizzare la dispersione degli elettroni. I fotoni entrano in un intensificatore a focalizzazione elettrostatica *(diagramma in basso)* attraverso una spessa finestra di fibre ottiche che sigilla la parte interna del tubo intensificatore ove è fatto il vuoto. Il fotocatodo si trova sulla faccia interna della finestra; gli elettroni che esso emette sono focalizzati da un campo elettrico sulla piastra a microcanali. In questo modo l'immagine intensificata risulta capovolta e quindi il tubo intensificatore può essere facilmente accoppiato con altri sistemi ottici, che capovolgono anch'essi l'immagine. L'immagine di uscita risulta così diritta. Cambiando il fuoco del fascio di elettroni l'immagine può anche essere ingrandita o ridotta.

Anodi di forme diverse registrano la posizione di un impulso di elettroni che emerge da una piastra a microcanali in modo tale da consentire l'immagazzinamento e il trattamento dei dati dell'immagine. La posizione degli impulsi rivelata da una schiera di anodi separati offre uno scarso potere risolutivo *(a)*. Per un potere risolutivo più elevato il numero di anodi richiesto rende i circuiti proibitivamente complessi e si preferisce quindi adottare soluzioni differenti. Alcune di queste stabiliscono la posizione di un impulso, depositando la carica su una coppia di fili facenti parte di due griglie, una orizzontale e una verticale. Ogni filo della griglia può essere dotato di un circuito di rilevamento *(b)*, oppure si può codificare la posizione dell'impulso sotto forma di numero binario *(c)*. L'ultima configurazione richiede un minor numero di circuiti di rilevamento a parità di potere risolutivo. Altre soluzioni rilevano una nube di carica su varie porzioni dell'anodo o determinano la posizione dell'impulso confrontando la carica in un circuito con quella in un altro circuito. È necessario calcolare i rapporti di carica in direzione verticale e orizzontale per determinare le coordinate spaziali dell'impulso. La carica può essere misurata a entrambe le estremità di due fili capacitivi *(d)*, ai quattro angoli di una piastra resistiva *(e)*, sui quattro quadranti di un disco *(f)* o su configurazioni geometriche più complesse di anodi isolati elettricamente *(g)*. In ogni caso i rapporti di carica rilevati dai vari circuiti possono dare la posizione di un impulso sulla piastra a microcanali con la precisione di alcuni millesimi del diametro del campo visivo.

verso il fotocatodo è impedito dalle collisioni contro le pareti. Sebbene il modo più semplice di realizzare canali del genere sia per incisione fotolitografica, il processo, come dianzi precisato, non offre risultati soddisfacenti dal punto di vista tecnico. Nelle applicazioni ove è richiesto un guadagno elevato due o più piastre vengono montate una ruotata rispetto all'altra.

Poiché le facce del wafer non sono tagliate ad angolo retto rispetto ai fori dei canali, ogni microcanale risultante presenta un angolo lungo il percorso. Le successive cascate di elettroni seguono facilmente il percorso angolato, mentre ciò non è possibile per gli ioni perché non danno origine a cascate di altri ioni. Mediante questo artificio la tensione applicata (e di conseguenza il guadagno) può essere notevolmente aumentata prima che si raggiunga la soglia della retroazione ionica rigenerativa.

Perseguendo la ricerca per realizzare microcanali curvi con cui minimizzare la retroazione ionica, gruppi di ricercatori in vari laboratori cercano di sfruttare la variazione prevedibile della viscosità del vetro con la temperatura. Se una piastra a microcanali di vetro viene serrata fra due blocchi metallici tenuti a temperature diverse, nella piastra si forma un gradiente uniforme di temperatura. In prossimità del blocco più caldo il vetro è più fluido. Se la piastra viene «stirata» mediante spostamento laterale di uno dei blocchi, i microcanali si piegano di più in prossimità del blocco più caldo, che non in prossimità di quello più freddo. In questo modo ogni microcanale assume una forma curva. Anche se le piastre a canali curvi sono ancora in fase sperimentale, possono moltiplicare gli elettroni secondari per un fattore di circa un milione senza retroazione ionica e presentano una risoluzione delle caratteristiche spaziali pari a quella che si ottiene mediante le piastre a microcanali convenzionali.

Una terza limitazione dei canali moltiplicatori è la densità di carica degli elettroni nel canale, chiamata carica spaziale. Quando la cascata di elettroni raggiunge una densità lineare di circa 10 milioni di elettroni per millimetro, la repulsione elettrostatica reciproca tende a riportare gli elettroni secondari aggiuntivi sulla superficie del canale prima che il campo possa imprimere loro un'accelerazione significativa. Ne deriva un'ulteriore limitazione nella crescita della cascata. In molte applicazioni quest'effetto è desiderabile in quanto rende gli impulsi di uscita dal moltiplicatore press'a poco identici; le fluttuazioni casuali nella carica complessiva portata da ogni cascata vengono di conseguenza eliminate.

La quarta limitazione della piastra a microcanali ne controlla il rendimento. La superficie complessiva di ingresso dei canali è inferiore alla superficie dell'intera piastra. Nel caso di canali circolari, il rapporto fra la sezione trasversale complessiva dei canali e quella della piastra deve essere inferiore a 0,91 per limitazioni di tipo geometrico; a causa dello spessore delle pareti dei canali il rapporto predetto scende in molti casi ad appena 0,55. All'incirca metà del flusso di ingresso colpisce, quindi, le zone chiuse rivestite di metallo, comprese fra gli ingressi dei canali.

È possibile adottare alcuni accorgimenti per ridurre questa perdita. Se alla parte frontale della piastra è applicato un campo elettrostatico di forte intensità, gli elettroni emessi dalle zone chiuse possono essere trascinati verso l'ingresso di un canale adiacente e avviare così una cascata. È possibile anche impartire agli ingressi dei canali una forma svasata mediante attacco chimico, elevando così il rapporto fra superficie d'ingresso dei canali e superficie totale della piastra. Diversi laboratori stanno eseguendo esperimenti con canali quadrati o esagonali, che possono essere assiemati in modo da offrire una superficie chiusa complessiva minore di quanto si può ottenere con canali a sezione trasversale circolare.

Non è necessario che l'uscita di una piastra a microcanali sia raccolta su uno schermo fluorescente. In molte applicazioni, come per esempio nel telerilevamento di dati fotometrici in un veicolo spaziale, interessa avere un'uscita elettrica e non un'uscita ottica. Il più semplice di tali sensori è il fotomoltiplicatore a microcanali; questo assomiglia a un intensificatore a microcanali con focalizzazione di prossimità nel quale lo schermo fluorescente sia stato sostituito da un semplice anodo metallico. L'anodo fornisce un impulso elettrico di uscita ogni volta che la piastra a microcanali emette una cascata di elettroni. Grazie al ritardo estremamente breve fra lo stimolo e l'impulso amplificato di uscita, il fotomoltiplicatore a microcanali è stato impiegato per il conteggio di fotoni singoli e anche per effettuare misurazioni del tempo di transito di particelle subatomiche prodotte negli acceleratori.

Un fotomoltiplicatore convenzionale possiede un solo collegamento per l'uscita elettrica e non è quindi adatto per il rilevamento di immagini. Un fotomoltiplicatore a microcanali può, d'altro canto, essere progettato in modo che la posizione di un impulso di uscita sull'anodo possa essere decodificata elettricamente. Dal punto di vista concettuale il modo più semplice per ottenere questa prestazione è quello di dotare ogni pixel di un anodo e di un circuito di rilevamento. La tecnica fornisce solo immagini alquanto scadenti poiché il numero dei collegamenti anodici e dei circuiti di scatto (che deve essere uguale al numero dei pixel) non può essere aumentato indefinitamente.

Per ridurre il numero dei circuiti di scatto ogni impulso può essere rilevato simultaneamente su due griglie sovrapposte di fili verticali e orizzontali. Se l'impulso fa scattare il filo verticale e quello orizzontale più vicini delle griglie, il numero dei circuiti richiesti varia con la radice quadrata del numero dei pixel. Disposizioni più complesse possono dare codificazioni ancora più efficienti dell'informazione di posizione. Se, in particolare, l'impulso può far scattare qualsiasi combinazione di fili nelle griglie, una qualsiasi combinazione possibile di fili può codificare una corrispondente posizione dell'impulso. Con una configurazione del genere il numero dei circuiti varia solo con il logaritmo del numero dei pixel.

Un altro tipo di anodo sensibile alla posizione codifica le coordinate dell'impulso dell'elettrone in modo continuo e analogico. Una disposizione sviluppata di recente impiega solo tre circuiti rivelatori. L'anodo è costituito da tre elettrodi metallici situati a notevole distanza dalla faccia di uscita del wafer a microcanali, di modo che ciascuno degli elettrodi intercetta una porzione della nube di elettroni associata a ogni cascata di elettroni. Uno degli elettrodi ha una forma a zig-zag e divide il piano dell'anodo in due parti *(si veda l'illustrazione della pagina a fronte)*. Da una parte, l'area dell'elettrodo collettore delle cariche varia linearmente con la coordinata x ed è indipendente dalla coordinata y. Dall'altra parte dell'anodo a zig-zag l'area dell'elettrodo varia linearmente con la coordinata y ed è indipendente dalla coordinata x. Se si amplificano e si misurano gli impulsi di carica ricevuti dai singoli elettrodi in modo indipendente uno dall'altro e si calcolano i rapporti fra gli impulsi dipendenti da x e da y e la somma dei tre impulsi, si può stabilire con precisione la posizione dell'impulso di uscita. Questo tipo di anodo è stato progettato originariamente da Hal O. Anger della Università della California a Berkeley.

Insieme ai miei colleghi dello Space Sciences Laboratory di Berkeley ho adottato lo schema di Anger per due applicazioni destinate a veicoli spaziali. Altri quattro gruppi stanno cercando di incorporare il rivelatore in fotomoltiplicatori per il conteggio di fotoni alle lunghezze d'onda del visibile. Tali rivelatori sensibili alla posizione possono funzionare da dispositivi di ingresso per un calcolatore che può registrare e analizzare i dati. La piastra a microcanali può quindi fungere da intermediario fra radiazione incidente e dati di uscita di varietà quasi illimitata. Sistemi di questo tipo possono condurre a progressi significativi in fisica e in astronomia.

Gli anelli nel sistema solare

Non solo Saturno, ma anche Giove e Urano possiedono anelli. Le recenti missioni spaziali hanno dimostrato che questi anelli sono costituiti in realtà da una miriade di sottoanelli tenuti insieme da forze interagenti

di James B. Pollack e Jeffrey N. Cuzzi

L'avvicinarsi della sonda spaziale *Voyager 2* a circa 100 000 chilometri da Saturno, il 25 agosto 1981, contraddistingue il clima di un periodo di esplorazione planetaria nel quale gli anelli di Saturno ci hanno riservato sorprese forse paragonabili a quelle avute dai primi uomini che li videro quasi quattro secoli fa. Al momento attuale gli anelli di Saturno sono conosciuti in una serie di dettagli, quali le bande, le razze e le trecce. Alcuni dettagli sono tuttora non spiegati. D'altro lato risulta che non solo Saturno è circondato da anelli. Giove possiede un anello e Urano ne possiede almeno nove distinti. Resta ancora da stabilire se Nettuno abbia anelli e se gli anelli siano quindi comuni ai pianeti gassosi giganti del sistema solare esterno. Esporremo qui la struttura e la composizione degli anelli di Giove, Saturno e Urano, con particolare riguardo a quelli di Saturno. Esporremo poi quanto è noto sui processi che determinano la formazione degli anelli. Infine prenderemo in considerazione alcune spiegazioni alternative di come si formano gli anelli.

Strane appendici

Gli anelli di Saturno furono osservati per la prima volta nel luglio del 1610 da Galileo Galilei. In parte per la povertà delle immagini prodotte dalla sua invenzione, il telescopio astronomico, in parte per aver egli scoperto i quattro satelliti maggiori di Giove solo pochi mesi prima, Galileo ritenne che le strutture confuse, a forma di orecchie, fossero due satelliti prossimi a Saturno. Presto la sua opinione mutò, in seguito al fatto che le strane appendici non variavano di giorno in giorno la propria posizione rispetto a Saturno; inoltre, nel 1612, le appendici scomparvero. Oggi sappiamo che in quell'anno gli anelli si erano orientati di taglio verso la Terra, rendendosi quindi quasi totalmente invisibili.

Per gli astronomi la geometria delle appendici rappresentò un quesito di difficile soluzione. Si suppose che tali appendici fossero manici attaccati a Saturno, oppure che fossero l'insieme di parecchi satelliti in orbita solo intorno alla parte posteriore di Saturno, in modo da non produrre mai ombra sul pianeta. Infine, nel 1655, Christiaan Huygens prospettò l'ipotesi che le appendici fossero il segno visibile di un disco sottile e piatto di materia distaccato dal pianeta e disposto sul suo piano equatoriale. A seconda della posizione di Saturno e di quella della Terra nelle loro orbite intorno al Sole il disco varierebbe la propria inclinazione verso la Terra; ne conseguirebbe un'immagine variante da una linea sottile a un'ampia ellisse.

Nei due secoli che seguirono si credette che il disco fosse una lamina continua di materia. Una difficoltà alla conferma di questa ipotesi emerse nel 1675, non appena Jean Dominique Cassini scoprì che la banda scura, ora nota come divisione di Cassini, separa il disco in due anelli concentrici. Inoltre, verso la fine del diciottesimo secolo, Pierre Simon de Laplace dimostrò che, se gli anelli fossero stati una vasta e compatta lamina di materia, le forze combinate dovute alla gravità di Saturno e alla rotazione del disco l'avrebbero lacerata. È fondamentale che ogni particella del disco conservi la propria distanza radiale da Saturno perché due forze sono in equilibrio. La forza di gravità è centripeta e attira la particella, l'altra è centrifuga e la spinge lontano. La forza centrifuga è conseguenza della velocità di rotazione, quindi il disco ruota. Il problema consiste nel fatto che, nel caso di un disco rigido rotante, le forze sono in equilibrio solamente a una data distanza radiale. Laplace quindi avanzò l'ipotesi che gli anelli di Saturno fossero formati da tanti sottoanelli stretti, abbastanza sottili da sopportare il lieve squilibrio di forze lungo la propria larghezza radiale.

Il passo finale verso la concezione moderna degli anelli fu compiuto nel 1857, quando James Clerk Maxwell ottenne il premio Adams dall'Università di Cambridge per la dimostrazione matematica che i sottoanelli sono formati da numerose minuscole masse, in orbite indipendenti. Dimostrazioni sperimentali a favore di questa ipotesi vennero da molte parti. Nel 1895, per esempio, gli astronomi americani James E. Keeler e William W. Campbell dedussero le velocità delle particelle negli anelli in base allo spostamento Doppler, cioè in base alle variazioni di frequenza delle righe dello spettro della luce solare riflessa sulla Terra da tali particelle. Essi scoprirono che la velocità di rotazione degli anelli intorno a Saturno è diversa da quella dell'atmosfera del pianeta. Inoltre le parti interne degli anelli ruotano a una velocità superiore a quella delle parti esterne, esattamente in accordo con le leggi della fisica relative a particelle rotanti in orbite indipendenti.

Gli anelli di Urano furono scoperti per caso. Alcuni gruppi di astronomi si erano proposti di osservare l'occultazione della stella SAO 158687 da parte di Urano il 10 marzo 1977, con lo scopo di studiare la

L'anello «*A*» di Saturno è stato fotografato dalla sonda spaziale *Voyager 2* il 23 agosto 1981. Nell'immagine risultante, in falsi colori, l'anello *A* è giallastro; l'intervallo più esterno nell'anello è la divisione di Encke. La parte dell'anello all'esterno della divisione di Encke mostra una serie di bande poco luminose, più fitte vicino all'orbita di 1980S27, un satellite scoperto nelle immagini riprese da *Voyager 1* e visibile in alto a sinistra come un punto bianco. Si pensa che le bande siano l'effetto di risonanze nell'anello, dovute agli effetti gravitazionali del satellite. Si ritiene inoltre che 1980S27 e un altro satellite, 1980S26, diano forma all'anello *F* di Saturno, troppo debole per essere qui visibile. I sottoanelli biancastri e bluastri in basso a destra occupano la divisione di Cassini. La loro differente colorazione indica la diversa composizione delle loro particelle.

GIOVE

SATURNO

URANO

RAGGI PLANETARI

DISCO DIFFUSO

ANELLO LUMINOSO

1979J3?

1979J1

AMALTHEA

1979J2

DIVISIONE DI CASSINI

DIVISIONE DI ENCKE

1980S28

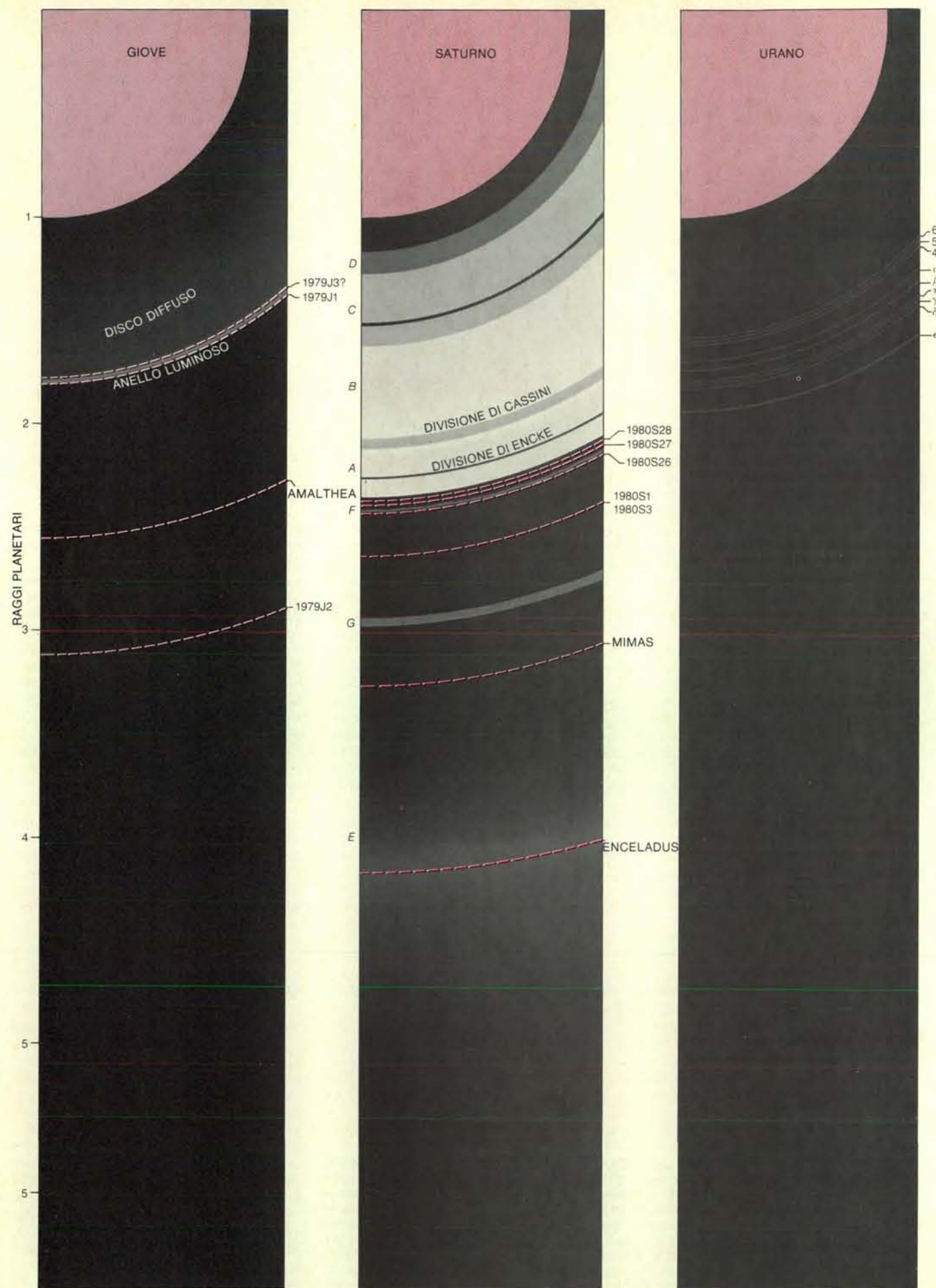
1980S27

1980S26

1980S1

1980S3

MIMAS

ENCELADUS

struttura dell'atmosfera del pianeta. Le osservazioni che ottennero i risultati migliori furono quelle effettuate da James L. Elliot e dai suoi collaboratori a bordo del Kuiper Airborne Observatory, un veicolo equipaggiato con un telescopio da 91 centimetri. Il suo gruppo (e alcuni altri) scoprirono che la luminosità della stella diminuiva non solo quando passava dietro Urano, ma anche in un certo numero di posizioni prossime al pianeta, anche ben al di fuori della sua atmosfera. Le rapide variazioni di luminosità definivano una serie di distanze da un lato di Urano grosso modo simmetrica a quella dalla parte opposta del pianeta. La simmetria fu attribuita alla presenza di anelli quasi circolari, molto stretti e abbastanza opachi. Ulteriori osservazioni hanno rilevato sinora la presenza di nove anelli, tutti giacenti entro un raggio planetario dal limite superiore dell'atmosfera di Urano.

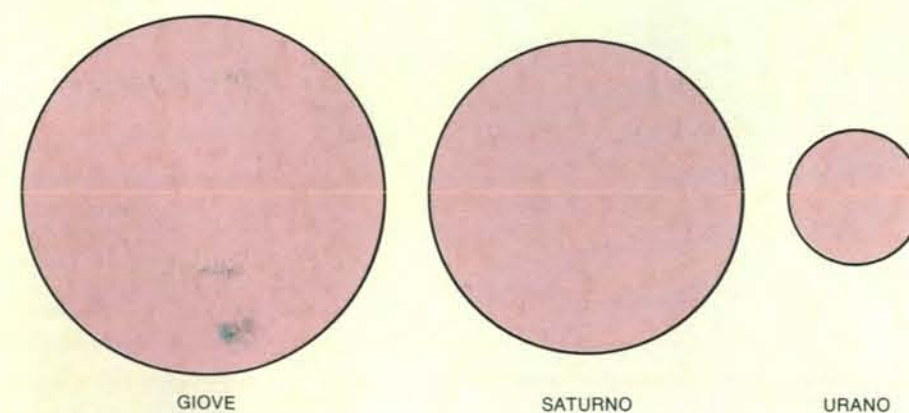
Gli anelli di Giove sono stati scoperti da sonde spaziali. Il primo indizio della loro presenza si ebbe quando *Pioneer 10* passò vicino a Giove nel 1974. Questo pianeta possiede un campo magnetico che intrappola le particelle cariche in regioni dello spazio circostante. Tali regioni intorno a Giove corrispondono alle fasce di Van Allen della Terra. *Pioneer 10* registrò una diminuzione del numero di particelle ad alta energia nelle fasce intorno a Giove fra 50 000 e 55 000 chilometri al di sopra dell'atmosfera del pianeta. Mario H. Acuña e Norman F. Ness del Goddard Space Flight Center della National Aeronautics and Space Administration suggerirono l'ipotesi che detta diminuzione potesse essere attribuita al parziale assorbimento delle particelle da un satellite vicino o da un sistema di anelli. Fu provato che quest'ultima ipotesi corrispondeva al caso in esame. Il debole anello di Giove fu fotografato nel 1979 da *Voyager 1* e *Voyager 2*. Le traiettorie dei veicoli spaziali furono deviate verso Saturno dal campo gravitazionale di Giove. Essi giunsero in prossimità di Saturno nel novembre del 1980 e nell'agosto del 1981.

Le caratteristiche degli anelli

Gli anelli di Saturno, Urano e Giove presentano un certo numero di caratteristiche comuni. Innanzitutto essi sono formati da miriadi di particelle su orbite indipendenti. In secondo luogo essi sono molto più vicini al pianeta di quanto non lo siano i satelliti maggiori; infatti la massa di ogni sistema di anelli dista dalla superficie del pianeta meno di un raggio del pianeta stesso. Inoltre gli anelli giacciono sul piano equatoriale del pianeta: in realtà quasi tutta la materia che forma gli anelli è contenuta in una sottile regione su quel piano. Infine i sistemi di anelli di Giove e di Saturno hanno un certo numero di minuscoli satellitini prossimi o all'interno degli anelli. Si suppone che simili satellitini si trovino anche negli anelli di Urano. Eppure ogni sistema di anelli presenta proprie caratteristiche peculiari. Il sistema di Saturno presenta sette sezioni principali. Alcune di queste sono separate dalle sezioni vicine da intervalli anulari più o meno vuoti; i confini delle altre sono caratterizzati da variazioni di concentrazione di particelle nell'anello. Le sezioni sono contraddistinte da lettere: la successione alfabetica non è legata alle distanze da Saturno, ma all'ordine in cui tali sezioni sono state scoperte o ipotizzate.

Il corpo principale del sistema di anelli di Saturno include gli anelli *A* e *B*, luminosi e relativamente opachi. Essi sono separati dalla divisione di Cassini larga 5000 chilometri, che appare piuttosto trasparente, ma non vuota. Il corpo principale del sistema di anelli di Saturno include anche l'anello *C*, più debole, meno opaco, situato entro il bordo interno dell'anello *B*. L'anello *C* ha un grado di opacità paragonabile a quello della divisione di Cassini. L'anello *D*, ancora più debole, è situato all'interno dell'anello *C*. Infine, tre debolissimi anelli *(E, F* e *G)* sono situati all'esterno dell'anello *A*. Presi insieme, gli anelli principali di Saturno *(A, B* e *C)* misurano circa 275 000 chilometri di diametro, corrispondenti a circa tre quarti della distanza tra la Luna e la Terra. In confronto lo spessore degli anelli di Saturno è trascurabile. Un limite superiore di circa un chilometro è stato attribuito alla loro estensione in senso verticale. In confronto alla loro ampiezza gli anelli sono migliaia di volte più sottili di una lama da rasoio.

Informazioni sulla composizione delle particelle degli anelli di Saturno si possono ottenere dal modo in cui gli anelli stessi riflettono o assorbono luce di differente lunghezza d'onda. Per esempio, gli anelli *A, B* e *C* riflettono scarsamente la luce del Sole in certe lunghezze d'onda dell'infrarosso vicino. Questa proprietà, caratteristica del ghiaccio d'acqua, lascia supporre che il principale componente della superficie delle particelle che formano questi anelli sia ghiaccio. Quest'ultimo, però, è di colore bianco, ossia riflette più o meno uniformemente tutte le lunghezze d'onda visibili. Le particelle degli anelli *A, B* e *C* riflettono invece meno la luce blu della rossa. Forse sono presenti in piccole quantità altre sostanze. Come causa del colore rossastro, si è pensato a polvere contenente ossido di ferro. È stato anche ipotizzato che a generare il colore rossastro siano i composti prodotti dalla radiazione ultravioletta del Sole. Alcuni composti incolori contenenti zolfo, se sottoposti a radiazione ultravioletta, si trasformano in polisolfuri, che hanno la proprietà di assorbire selettivamente la radiazione blu. Dai dati trasmessi dai Voyager è stato rilevato, con sorpresa, che le particelle degli anelli *A* e *B* presentano colore simile, ma sono più brillanti e più rosse di quelle dell'anello *C* e della divisione di Cassini.

Osservazioni a mezzo radar forniscono ulteriori informazioni. Nel 1973 Richard M. Goldstein e Gregory Morris del Jet Propulsion Laboratory del California Institute of Technology sondarono gli anelli di Saturno con onde radar di cui rilevarono gli echi per mezzo di un'antenna di 64 metri della rete di rilevamento, situata a Goldstone, in California. L'alto potere riflettente degli anelli *A* e *B* evidenziò che la maggior parte delle particelle in quei due anelli sono paragonabili come dimensioni alla lunghezza d'onda radar di alcuni centimetri impiegata dagli osservatori. Se fossero state più piccole della lunghezza d'onda sarebbero state trasparenti alle onde radar. Se fossero state molto più grandi della lunghezza d'onda, l'emissione di radiazioni termiche a quelle lunghezze d'onda sarebbe stata significativa. Il basso livello di queste radiazioni limita la loro

I sistemi di anelli intorno a Giove, Saturno e Urano sono tracciati in scale in cui il raggio di ciascun pianeta vale 1 unità. Giove *(a sinistra)* possiede un anello luminoso che è molto debole e quasi trasparente. Un disco di particelle ancora meno luminoso si estende dall'anello verso l'interno, forse fin dentro l'atmosfera del pianeta. Un alone di particelle (non indicato) dà al sistema uno spessore verticale di circa 20 000 chilometri. Il sistema fu scoperto da una sonda spaziale. Saturno *(al centro)* possiede sette anelli. Gli anelli *A* e *B* sono separati dalla divisione di Cassini; l'anello *A* comprende la divisione di Encke. Soltanto gli anelli principali *(A, B* e *C)* sono visibili per mezzo di telescopi situati sulla Terra; gli altri, eccetto l'anello *E*, sono stati scoperti da sonde. Urano *(a destra)* possiede non meno di nove anelli. Qui la loro larghezza è stata ingrandita. Questi sono stati rilevati da terra e sono indicati con numeri o lettere greche. La denominazione di satelliti nei sistemi di anelli è riportata alla destra di ogni riquadro; le orbite sono raffigurate con linee tratteggiate. Le dimensioni relative di Giove, Saturno e Urano sono mostrate qui sopra.

Le razze e le bande negli anelli di Saturno sono state fotografate per la prima volta da *Voyager 1*. Le razze si formano in modo sporadico nell'anello *B*; sono cunei disposti più o meno radialmente. Ciascuna razza appare luminosa nelle immagini della faccia degli anelli illuminata dal Sole, riprese quando la sonda e il Sole si trovano da parti opposte rispetto a Saturno, in modo che la luce solare si riflette in avanti, dalle particelle dell'anello alla sonda *(fotografia in alto)*. Al contrario le razze appaiono scure quando la sonda e il Sole si trovano dalla stessa parte rispetto a Saturno, e, quindi, la luce del Sole viene riflessa indietro *(fotografia in basso)*. In ambedue le immagini risulta evidente la suddivisione degli anelli in sottoanelli. L'ampia banda scura è la divisione di Cassini.

dimensione a non più di qualche metro.

I dati trasmessi dal Voyager hanno confermato ed esteso le precedenti scoperte. In un esperimento sono state inviate dalla sonda spaziale verso la Terra onde radio attraverso gli anelli e fu misurata la quantità di energia diffusa dalle particelle degli anelli a vari angoli di deviazione rispetto all'originale percorso delle onde. Con l'aumentare delle dimensioni delle particelle rispetto alla lunghezza d'onda, la figura di diffusione diviene sempre più concentrata entro piccoli angoli di deviazione in avanti. Dall'analisi dei dati del Voyager, G. Leonard Tyler e Ahmed Essam A. Marouf, della Stanford University, trassero l'indicazione che le particelle di maggiori dimensioni negli anelli *A*, *B* e *C* hanno un diametro di circa 10 metri, mentre le più numerose sono di circa 10 centimetri, con differenze da zona a zona nella distribuzione delle dimensioni.

Come mediante la misurazione della diffusione di onde radar è possibile rilevare particelle dell'ordine di grandezza di una lunghezza d'onda radar, così, mediante la diffusione di luce solare, è possibile rilevare particelle della dimensione della lunghezza d'onda della luce visibile. In particolare il forte illuminamento di una parte di un anello, quando è visto sotto piccoli angoli di diffusione in avanti, implica l'esistenza di numerose particelle di dimensioni dell'ordine di un micrometro nella parte stessa dell'anello. Tale osservazione può essere realizzata solo quando Saturno si trova fra il Sole e l'osservatore. È ovvio che simili osservazioni possono essere effettuate solo da veicoli spaziali e non dalla Terra, dove tale condizione non si verifica mai. Gli studi sui dati del Voyager evidenziano pertanto che le particelle di dimensioni intorno al micrometro costituiscono una grossa quota delle particelle presenti nell'anello *F*, una quota significativa in molte parti dell'anello *B* e una quota meno rilevante nella parte esterna dell'anello *A*. D'altro canto, l'anello *C* e la divisione di Cassini non recano traccia di tali particelle di piccole dimensioni.

Strutture negli anelli di Saturno

Prima del passaggio della sonda spaziale Voyager nei pressi di Saturno ben poco era noto delle strutture degli anelli del pianeta. La divisione di Cassini era certo conosciuta, come pure la divisione di Encke, una banda più stretta nella parte esterna dell'anello *A*. Le fotografie ad alto potere risolutivo degli anelli, eseguite dal Voyager, riservarono un certo numero di sorprese. Sono apparse strette regioni anulari di differente luminosità e opacità, numerose quanto i solchi di un disco fonografico. Sono state inoltre trovate strutture particolari che non hanno andamento circolare. Ne sono esempi i cunei disposti radialmente, come razze di una ruota, nell'anello *B* e i nodi, le trecce e gli attorcigliamenti dell'anello *F*.

Il maggior numero di strutture particolari è presente nell'anello *B*, che esprime anche la maggiore densità di particelle rinvenuta negli anelli. Variazioni di opacità nell'anello *B* si hanno a piccole distanze radiali comprese tra 10 e 50 chilometri, mentre l'anello stesso ha una larghezza complessiva di 25 000 chilometri. La parte centrale, più opaca, dell'anello *B* è quella in cui sono presenti le razze. Generalmente, ogni razza può essere osservata per buona parte delle dieci ore occorrenti a una particella dell'anello *B* a compiere un'intera rivoluzione orbitale. Nel frattempo nuove razze sorgono sporadicamente in nuove parti dell'anello. Rispetto alle regioni circostanti, le razze appaiono brillanti in luce diffusa in avanti e scure in luce diffusa all'indietro. Da ciò si deduce l'abbondanza nelle razze di particelle di dimensioni dell'ordine del micrometro.

Ogni parte di una razza ruota intorno a Saturno alla stessa velocità delle particelle dell'anello che si trovano alla stessa distanza radiale. Le parti più interne hanno una velocità maggiore, perciò una razza si inclina sempre di più col tempo. Può capitare che scompaia. La stretta parte terminale di ogni razza cuneiforme sembra quasi coincidere con la distanza da Saturno alla quale il periodo di una particella orbitante coincide con il periodo di rotazione di Saturno. Il campo magnetico di Saturno è agganciato al pianeta e quindi ruota con esso. Di conseguenza le forze elettromagnetiche possono essere parzialmente responsabili della presenza delle razze. A questo proposito può essere importante il fatto che intense scariche elettrostatiche a larga banda siano state rilevate durante l'esperimento di radioastronomia planetaria condotto da *Voyager 1*. Sembra che tali scariche provengano da sorgenti nell'anello *B*, vicino alle regioni di intensa attività delle razze.

Rispetto all'anello *B*, l'anello *A* è privo di tratti caratteristici e uniformemente luminoso. Esso mostra, tuttavia, una varietà di stretti e distinti lineamenti nelle regioni più esterne. Innanzitutto la scura e relativamente vuota divisione di Encke, che ha una larghezza di circa 350 chilometri, contiene due o tre stretti sottoanelli irregolari, composti di minuscole particelle. Le immagini realizzate da *Voyager 2* mettono in evidenza, nei sottoanelli, dei cappi che ricordano quelli rilevati da *Voyager 1* nell'anello *F*. Quattro strane bande con struttura irregolare e ondulata racchiudono la divisione di Encke. A una maggiore distanza da Saturno, il bordo esterno dell'anello *A* evidenzia molte bande strette, larghe circa 20 chilometri, in cui abbondano piccole particelle. La spaziatura fra le bande decresce verso l'esterno. Il bordo esterno dell'anello *A* è tutto frastagliato. Anche qui c'è abbondanza di piccole particelle.

L'anello *C* e la divisione di Cassini hanno molte somiglianze strutturali, oltre al fatto di essere poco rossi, al loro grado di trasparenza simile e alla comune scarsità di particelle piccole. Ambedue mostrano bande discrete, regolarmente spaziate, di uniforme luminosità. Entrambe inoltre contengono una varietà di stretti intervalli a bordo netto, completamente vuoti, con una larghezza radiale che varia da 50 a 350 chilometri. Inoltre, alcuni di questi intervalli contengono sottoanelli a bordo netto, ancora più stretti, che sono completamente opachi. Molti di tali sottoanelli sono eccentrici (cioè non circolari) e non uniformi in larghezza. Le immagini pervenute da *Voyager 2* mostrano che i sottoanelli opachi spesso differiscono nel colore dal materiale circostante dell'anello. Essi sono più somiglianti alla materia degli anelli *A* e *B*.

La diffusione della radiazione elettromagnetica da parte delle particelle costituenti un anello permette di dedurre le dimensioni delle particelle presenti in abbondanza nell'anello stesso. In particolare, se una particella misura circa un decimo della lunghezza d'onda della radiazione incidente, essa diffonde la radiazione quasi uniformemente in tutte le direzioni *(a)*. Se è di poco più piccola della lunghezza d'onda, tende a diffondere la radiazione in avanti *(b)*. Se è più grande, diffonde la radiazione in avanti quasi per intero *(c)*. La lunghezza delle frecce rappresenta la quantità relativa di energia diffusa ai vari angoli. Osservando le particelle nelle razze dell'anello *B* di Saturno, che diffondono la luce del Sole per lo più in avanti, si deduce che le razze sono una concentrazione locale, temporanea, di particelle che misurano circa un micrometro.

Gli anelli di Urano e di Giove

Gli anelli di Urano sono sotto molti aspetti il complemento degli anelli di Saturno. Gli anelli principali di Saturno (*A*, *B* e *C*) sono ampi e inframmezzati da stretti intervalli. Gli anelli di Urano sono stretti (ciascuno misura radialmente da alcuni chilometri a 100 chilometri) e sono separati da larghe regioni vuote. Si è avuta conferma che gli anelli di Urano sono nove. Ognuno è stato denominato con un numero o una lettera greca. La tecnica dell'occultazione stellare utilizzata per studiare gli anelli consente, per la minuscola dimensione apparente della stella occultata, di ottenere una elevata risoluzione spaziale. Ciononostante molti anelli di Urano restano indistinti. Tali anelli hanno un'estensione in senso radiale inferiore ai cinque chilometri. (La tecnica dell'occultazione stellare, come è stata applicata nello studio degli anelli di Saturno da *Voyager 2*, consente di rilevare strutture dell'ordine di grandezza di alcune centinaia di metri.) Gli anelli più stretti di Urano (γ, δ e η) sono approssimativamente circolari e giacciono in uno stesso piano. Al contrario, gli anelli α, β e 4, un po' più lunghi, sono leggermente ellittici e inclinati rispetto al piano comune. L'anello ε è il più largo e il più ellittico dei nove. La sua larghezza radiale varia linearmente con la distanza da Urano; essa è di 20 chilometri nel punto più vicino al pianeta e di 100 chilometri nel punto più lontano. Sappiamo oggi che anche alcuni degli stretti sottoanelli eccentrici negli anelli di Saturno presentano una analoga variazione di larghezza.

L'indebolimento della luminosità delle stelle al loro passaggio dietro gli anelli di Urano indica che gli anelli hanno un'opacità paragonabile a quella delle parti più opache (e quindi più luminose) degli anelli di Saturno. L'anello ε è particolarmente opaco. Curiosamente esso è molto più opaco ai bordi. Le forze gravitazionali generate dalla forma di Urano, schiacciata ai poli, sono la causa del continuo cambiamento dell'orientamento degli anelli ellittici, ε incluso. L'anello ellittico compie una rotazione intorno al pianeta in circa

6300 ore. Durante quel tempo una particella nell'anello compie circa 750 orbite intorno a Urano. Due fattori influenzano dunque le caratteristiche dell'anello ε. Innanzitutto il suo orientamento cambia; inoltre le particelle dell'anello situate a differenti distanze radiali da Urano orbitano intorno al pianeta a velocità diverse, in accordo con le leggi fisiche dei moti orbitali. Nondimeno l'anello ε mantiene la propria integrità; persistono cioè la sua eccentricità, la sua larghezza variabile e la concentrazione delle particelle verso i bordi.

Di che cosa sono composti gli anelli di Urano? A causa della loro modesta larghezza essi sono difficili da osservare dalla Terra. Ciononostante, osservazioni recenti evidenziano che le particelle dell'anello sono molto scure, quindi il loro componente principale non è il ghiaccio. È molto verosimile che le particelle dell'anello orbitanti intorno a Urano siano silicati ricchi di composti che assorbono la luce del Sole. Alcuni ossidi di ferro e composti complessi di carbonio posseggono tali caratteristiche. Purtroppo nulla è noto sulla dimensione delle particelle.

Gli anelli di Giove sono costituiti da tre parti principali: un anello luminoso, un disco diffuso e un alone. L'anello luminoso ha una larghezza di circa 6000 chilometri. Il suo bordo esterno, nettamente definito, si trova a circa 58 000 chilometri (pari a 0,8 raggi planetari) al di sopra della superficie di Giove. Nella parte esterna dell'anello una stretta banda di circa 600 chilometri possiede una luminosità superiore di circa il 10 per cento a quella del resto. Inoltre l'opacità dell'anello «luminoso» è tanto bassa che solo lo 0,001 per cento della luce solare che lo attraversa è intercettato dalle sue particelle. Il disco diffuso è molte volte più debole. Esso si estende verso il pianeta dal bordo interno dell'anello luminoso. In realtà, non si esclude che possa anche arrivare fino all'atmosfera di Giove. Visti di taglio, l'anello luminoso e il disco presentano uno spessore che non supera i 30 chilometri. È notevole, comunque, il fatto che l'alone abbia uno spessore di circa 20 000 chilometri. L'alone è dunque molto più spesso del disco diffuso. Il suo bordo esterno giunge un po' oltre il bordo esterno dell'anello luminoso.

Studiando come l'anello luminoso brilla sotto piccoli angoli di diffusione, è stato scoperto che esso contiene particelle la cui dimensione caratteristica è di parecchi micrometri. Tali particelle sono del tutto inefficaci nell'assorbire protoni ed elettroni di alta energia. Da ciò si deduce che tali particelle non possono essere la causa della diminuzione del flusso di particelle di alta energia riscontrata nell'anello luminoso da *Pioneer 10*. Vi devono anche essere particelle di almeno un centimetro di diametro. Le particelle dell'anello sono rossastre, come molti asteroidi e satelliti del sistema solare esterno.

Collisioni negli anelli

L'architettura di un sistema di anelli dipende dall'interazione di molte forze. Fra queste figurano le forze gravitazionali, dovute ai satelliti esterni agli anelli e ai satellitini in essi incorporati; le forze elettromagnetiche, dovute al campo magnetico in rotazione del pianeta; inoltre le deboli forze esercitate dal mezzo gassoso diluito entro il quale ruotano gli anelli.

Tutte le particelle in un sistema di anelli sono animate da un comune moto orbitale intorno al pianeta: esse si muovono nello stesso senso di rotazione del pianeta. I moti verticali e radiali, sovrapposti al moto orbitale di ogni particella, non sono sottoposti a tale limitazione. Particelle vicine, muovendosi casualmente in queste direzioni l'una rispetto all'altra, provocano quindi collisioni. Quando le velocità relative casuali sono elevate, come potrebbero essere se gli anelli fossero una densa nube di particelle, le collisioni sono violente; anche se le collisioni sono molto rare, una grande quantità di energia delle particelle dovuta al moto relativo si trasforma e le riscalda deformandone la struttura. La conseguente perdita di energia causa la diminuzione rapida di queste velocità casuali. La diminuzione della componente verticale delle velocità porta a un appiattimento del sistema di anelli. La diminuzione della componente radiale porta a orbite quasi circolari. In breve, un anello grosso diventa molto presto un disco sottile e pressoché circolare.

Anche quando le particelle dell'anello hanno perduto quasi completamente il loro moto casuale, le collisioni continuano. La ragione è la seguente: la forza di gravità esercitata dal pianeta sulle particelle di un anello diminuisce con l'aumentare della distanza dal pianeta, cosicché le particelle più lontane dell'anello impiegano un tempo maggiore a compiere un'orbita intorno al pianeta. Quindi, la particella di un anello che percorre un'orbita di poco interna a quella di un'altra particella raggiunge quest'ultima e le due collidono se la loro distanza radiale è inferiore al diametro di una particella.

È verosimile che la velocità di collisione sia inferiore a un centimetro per secondo. Tuttavia ciò può trasformare una parte del moto circolare orbitale della particella in un moto verticale casuale. Collisioni successive eviteranno grossi incrementi di velocità verticale delle particelle e quindi si raggiungerà una situazione stazionaria che determinerà lo spessore dell'anello. Se le particelle hanno una vasta gamma di dimensioni, quelle più piccole acquisteranno velocità verticale soprattutto per l'attrazione gravitazionale esercitata su di esse da particelle più grosse che le sfiorano. Perderanno la loro velocità verticale principalmente per collisione con altre particelle di piccole dimensioni. In queste condizioni le particelle piccole raggiungeranno un'estensione verticale di molte volte superiore alla dimensione delle particelle abbondanti più grosse. Nel caso degli anelli *A* e *B* di Saturno le particelle piccole dovrebbero occupare uno spessore verticale compreso tra 10 e 100 metri. Le misurazioni eseguite dalle sonde Voyager rivelano che gli anelli principali di Saturno hanno uno spessore non superiore ad alcune centinaia di metri. Le misurazioni danno, per lo spessore osservato, valori compatibili con lo spessore previsto.

Per le collisioni, le particelle vicine trasformano una parte del loro moto circolare orbitale anche in moto radiale. Gli anelli, quindi, si espandono radialmente. Un anello isolato, senza limitazioni, si estenderà fino a quando le sue particelle si saranno allontanate a tal punto da non permettere ulteriori collisioni. Può darsi che l'anello luminoso di Giove abbia raggiunto questo stadio finale. La larghezza e la bassissima opacità di questo anello sono forse dovute alla dispersione delle sue particelle maggiori causata dalle collisioni verificatesi nel corso della vita del sistema solare.

Gli anelli di Saturno e Urano mostrano però contorni bruschi e netti che limitano regioni densamente affollate di particelle. Altri processi devono quindi contrastare la rapida diffusione dovuta alle frequenti collisioni. Una funzione importante, in questi processi, può essere svolta dai satellitini situati all'interno degli anelli o a essi adiacenti. I campi gravitazionali di satelliti più grandi e più lontani possono bloccare alcuni satellitini locali in orbite fisse e prevenire un loro spostamento dovuto alle interazioni gravitazionali con le particelle circostanti dell'anello.

Il moto di corpi in orbita intorno a un pianeta più massiccio è dominato essenzialmente dal campo gravitazionale del pianeta. In certi casi, però, l'attrazione gravitazionale fra due corpi orbitanti relativamente piccoli può essere amplificata in misura tale da consentire un'apprezzabile variazione di moto. Tale amplificazione è conosciuta come risonanza. Bisogna considerare che se il periodo orbitale di un satellite è un multiplo esatto o una frazione di quello di un altro satellite, l'effetto gravitazionale netto di ciascun satellite sull'altro è in sostanza una spinta o un'attrazione applicata ripetutamente nello stesso punto del moto ciclico. L'effetto risulta dunque amplificato. In alcuni casi la risonanza tiene legate coppie di satelliti in orbite i cui periodi stanno tra loro come due numeri interi piccoli. Tali moti in rapporto fisso tra loro sono ben noti nei sistemi di satelliti di Giove e Saturno, dove coinvolgono almeno quattro coppie di satelliti principali così come probabilmente varie coppie di satelliti di recente scoperta.

La risonanza e gli anelli

Una situazione molto diversa si ha nel caso di fenomeni di risonanza in un disco di particelle. In prossimità della distanza radiale dal pianeta in cui le particelle del disco avrebbero un periodo orbitale proporzionale a quello di uno dei satelliti del pianeta, l'amplificazione dell'effetto gravitazionale del pianeta provoca sul lungo periodo l'allontanamento delle orbite delle particelle dalla forma circolare. Così è probabile che le particelle siano indotte a collidere con quelle adiacenti meno perturbate e vengano allontanate dalla banda che si trova alla distanza radiale di risonanza. Tali bande hanno tipicamente una larghezza di alcune decine di chilometri e se ne hanno esempi negli anelli di Saturno tra cui le decine di stretti intervalli nella parte esterna dell'anello *A* che sembra siano il risultato di risonanze causate dai satelliti di recente scoperta denominati 1980S1, 1980S3, 1980S26 e 1980S27. La predominanza di particelle piccole vicino agli intervalli dimostra, con ogni probabilità, le violente collisioni provocate localmente da ogni risonanza.

Va ricordato anche un altro effetto delle risonanze che sono distanti da un piccolo satellite. Nel 1978 Peter M. Goldreich del Cal Tech e Scott D. Tremaine, attualmente all'Institute for Advanced Study di Princeton, avanzarono l'ipotesi che onde a spirale di fluttuazione nella densità della materia dell'anello fossero presenti negli anelli di Saturno. In precedenza era stata avanzata l'ipotesi che onde analoghe fossero responsabili della conformazione con bracci a spirale di galassie come la Via Lattea. Dalle immagini riprese da *Voyager 1* sembra che onde di densità a spirale siano debolmente visibili nella divisione di Cassini; la loro configurazione strettamente avvolta ricorda la molla di un orologio. Si ritiene che le onde vi insorgano per risonanza e può darsi che esse trasferiscano gli effetti delle risonanze a grandi distanze; sarebbero le onde quindi ad avere un ruolo fondamentale nel trasporto di materiale all'interno del disco formato dagli anelli di Saturno.

Le risonanze create in un disco da un satellite sono più vicine le une alle altre man mano che ci si avvicina all'orbita del satellite. A una certa distanza critica la spaziatura in senso radiale fra risonanze successive diventa uguale alla larghezza propria di ogni risonanza. All'interno di questa distanza critica le risonanze si sovrappongono. Come risultato si ha una zona continua di trasferimento di materia dell'anello, che la allontana dall'orbita del satellite. La larghezza della zona e il grado di evacuazione di materia dipendono dalla massa del satellite e dalla densità della materia circostante dell'anello.

Distorsioni negli anelli di Saturno risultano evidenti confrontando le immagini ottenute da *Voyager 2* ora dall'una ora dall'altra parte degli anelli. In ciascuna immagine l'anello *B* è la regione luminosa a sinistra; la regione scura a destra è la divisione di Cassini. Si nota che il bordo esterno dell'anello *B* non è circolare e si ha la conferma della presenza di un sottoanello eccentrico luminoso al di là del bordo dell'anello *B* (il sottoanello è stato scoperto da *Voyager 1*). Le particelle prossime al bordo esterno dell'anello *B* ruotano due volte intorno a Saturno, nel tempo impiegato dal satellite Mimas a percorrere un'orbita. La risonanza che ne deriva amplifica molto l'effetto gravitazionale di Mimas, provocando deformazioni nell'anello *B*, e provoca soprattutto rigonfiamenti ai lati opposti dell'anello. Al contrario, il sottoanello eccentrico ha un unico rigonfiamento che può essere il risultato del campo gravitazionale di un piccolo satellite non ancora individuato.

Le fotografie fatte da *Voyager 1* al suo passaggio presso Saturno nel novembre 1980 rivelarono trecce nell'anello *F*, ma in quelle fatte da *Voyager 2* nove mesi più tardi esse erano scomparse. L'anello stesso è a circa 80 000 chilometri (1,3 raggi di Saturno) dalla superficie del pianeta e a 4000 chilometri dal bordo esterno dell'anello *A*. Nel novembre 1980 *(in alto)* esso era costituito da tre filamenti, ciascuno largo 30 chilometri circa. I due filamenti esterni mostrano cappi, catene e nodi e sembra che si intreccino o per lo meno si intersechino. I campi gravitazionali dei satelliti 1980S26 e 1980S27 possono spiegare questa struttura distorta. Nell'agosto 1981 *(in basso)* l'anello *F* aveva mutato sostanzialmente la propria struttura. Un filamento non attorcigliato costituiva la caratteristica dominante, mentre gli erano associati tre filamenti di minore luminosità.

Maggiore è la massa del satellite, maggiore è la larghezza della zona, ma più è alta la densità di particelle nell'anello, più spesso accade che le collisioni fra particelle rimandino alcune di queste all'interno della zona. Jack Lissauer e Frank H. Shu, dell'Università della California a Berkeley e M. Hénon, dell'Osservatorio di Nizza, hanno dimostrato che se negli anelli di Saturno erano inglobati piccoli satelliti di diametro compreso fra alcuni chilometri e alcune decine di chilometri, i satellitini avrebbero potuto essere la causa di gran parte della struttura fine degli anelli. Da quanto si sa finora le immagini riprese da *Voyager 1* e *Voyager 2* non hanno però individuato la presenza di piccoli satelliti neanche nelle posizioni più verosimili.

Se il satellite è adiacente a un anello, può aversi anche un altro fenomeno: le risonanze sovrapposte che circondano l'orbita del satellite sono in grado di impedire l'espandersi dell'anello e di dargli un contorno netto. Il bordo esterno dell'anello *A* di Saturno è probabilmente mantenuto in questo stato dai satelliti 1890S26 e 1980S27. L'anello a sua volta respinge i satelliti. È il caso di 1980S26, molto prossimo alla distanza radiale da Saturno alla quale verrebbe intrappolato con il satellite molto più massiccio Mimas o Tethys. Tale risonanza «ancorerebbe» la sua distanza radiale in modo da consentirgli di continuare a modellare il bordo esterno dell'anello *A*. C'è un'unica difficoltà: in base a precise misurazioni, sembra che 1980S26 si trovi a una distanza leggermente diversa da quella di ancoraggio.

Infine, due satelliti le cui orbite siano prossime possono impedire la diffusione radiale di uno stretto anello tra esse compreso. Per esempio, 1980S26 e 1980S27, le cui orbite sono separate radialmente di soli 2000 chilometri circa, possono agire da «guardiani» per confinare lo stretto anello *F* a più filamenti nello spazio compreso fra le due orbite. Quando uno o entrambi i satelliti passano vicino a un dato filamento dell'anello, lo strappo gravitazionale può provocare distorsioni, come quelle che si osservano nelle immagini dell'anello riprese da *Voyager 1*. Non è chiaro, tuttavia, perché l'anello *F* debba essere formato da una molteplicità di stretti filamenti, e non da un filamento più grosso, e perché uno dei filamenti rilevati da *Voyager 1* non sia attorcigliato. Dalle immagini riprese da *Voyager 2* risulta che l'anello *F* non mostra la presenza di trecce o di altre distorsioni e sembra che la quantità dei filamenti sia mutata.

Il fatto che i satellitini che fungono da «guardiani» generino sottoanelli lisci (come gli stretti sottoanelli eccentrici dell'anello *C* di Saturno e nella divisione di Cassini e probabilmente nell'anello ε di Urano) o sottoanelli attorcigliati (come quelli dell'anello *F* di Saturno e quelli della divisione di Encke, il cui attorcigliamento fu rilevato grazie alle immagini riprese da *Voyager 2*), può dipendere dal grado di capacità dei sottoanelli oppure, ed è la stessa cosa, dalla densità delle particelle. In un sottoanello piuttosto trasparente le collisioni fra particelle sono piuttosto rare e, per questo motivo, gli «echi» delle perturbazioni passate possono persistere. In un sottoanello pressoché opaco la maggiore frequenza delle collisioni può smorzare le perturbazioni.

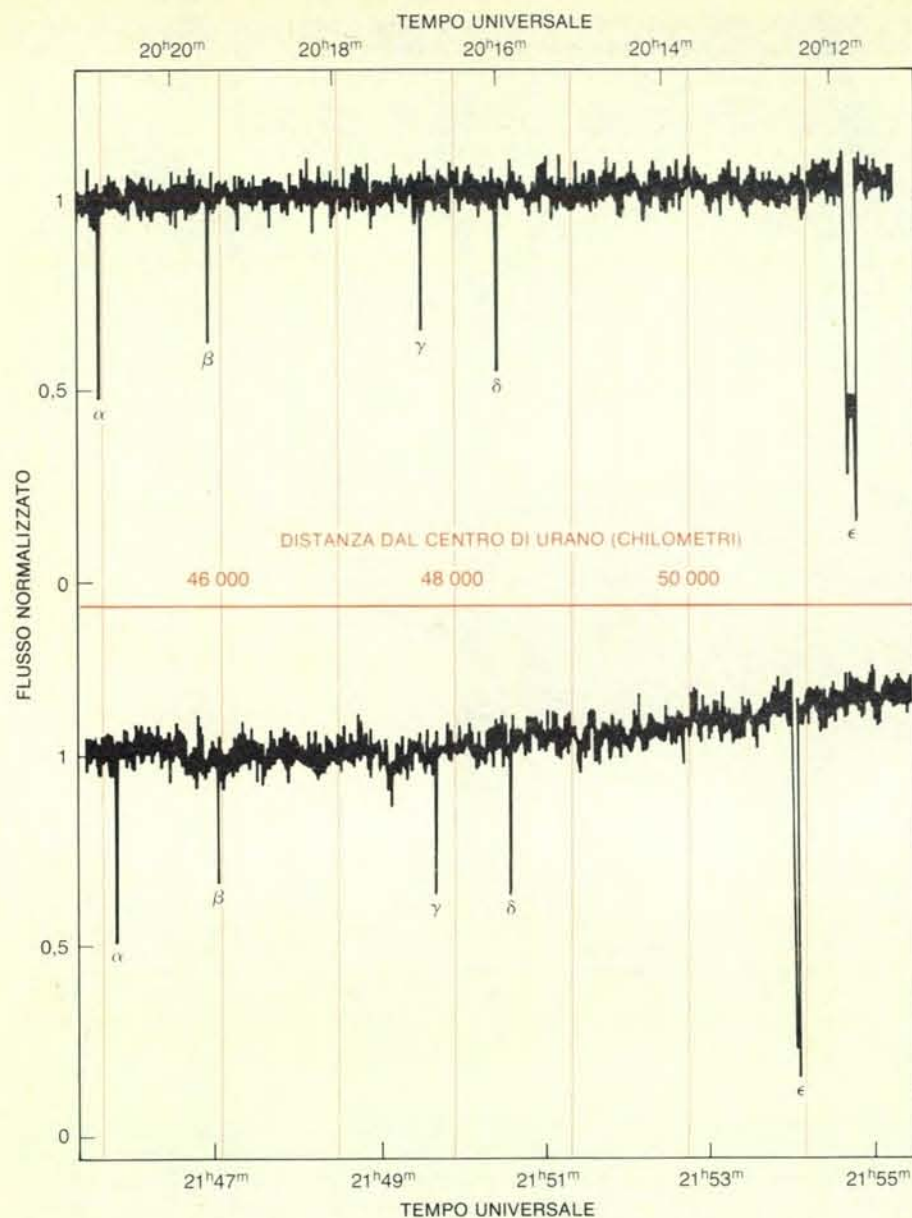
La funzione di sorveglianza dei confini entro cui si muovono le particelle, dovuta alle forze gravitazionali di satellitini all'interno di un sistema di anelli, spiega il perché di certe suddivisioni in bande degli anelli di Saturno. Il disegno in alto mostra un satellitino e (a titolo d'esempio) due particelle, in orbita intorno al pianeta. In accordo con le leggi fisiche, la particella interna si muove più rapidamente del piccolo satellite e questo, a sua volta, più rapidamente della particella esterna *(frecce in nero)*. Avviene che il satellitino venga superato dalla particella interna nello stesso momento in cui questa supera a sua volta la particella esterna *(frecce in colore)*. Ciascuna particella è attratta per gravitazione dal satellitino; ne consegue che, per il moto orbitale intorno al pianeta, ciascuna particella, subito dopo essersi trovata faccia a faccia con esso, si troverà più vicina di prima al satellitino. Il disegno in basso mostra il risultato. La pura attrazione gravitazionale esercitata sulla particella esterna è orientata nella direzione del moto orbitale della particella stessa, e, quindi, questa raggiunge un'orbita più alta. Quella esercitata sulla particella interna è invece orientata in senso contrario alla direzione del moto orbitale e la particella scende su un'orbita più bassa. Alla fine il satellitino ha spazzato intorno alla sua traiettoria una banda tanto più larga quanto maggiore è la sua massa.

Differenti gradi di trasparenza e quindi di densità di particelle negli anelli di Saturno risultano evidenti in una immagine ripresa da *Voyager 1*. Essa mostra la faccia non illuminata degli anelli, quando incrociano il disco luminoso del pianeta. Il disco brilla attraverso la maggior parte dell'anello *C (in basso)* e anche attraverso la divisione di Cassini, benché ambedue includano sottoanelli di capacità maggiore. Al contrario, gran parte dell'anello *B (al centro)* è quasi opaca.

Una piena comprensione della fisica delle interazioni fra anelli e satellitini non è ancora a portata di mano. Per esempio, Stanley F. Dermott e Thomas Gold della Cornell University hanno presentato un'ipotesi secondo la quale un sottoanello stretto ed eccentrico, con bordi netti, è mantenuto tale da un singolo piccolo satellite nascosto al suo interno e non da una coppia di satelliti o di satellitini. Il fatto di non aver trovato finora satellitini inglobati negli anelli principali di Saturno significa però che si deve esplorare in altre direzioni per spiegare la fine struttura radiale in regioni quali l'anello *B*. Un'ipotesi prospetta che le particelle dell'anello possano rispondere ai momentanei aumenti e diminuzioni della loro densità con successivi aumenti e diminuzioni. Supponiamo che la densità delle particelle aumenti in un certo punto. Le collisioni fra particelle diventano più numerose e dal momento che una parte dell'energia delle collisioni è perduta nel riscaldamento e nella deformazione delle particelle, l'energia del moto casuale diminuisce. Nonostante che le successive collisioni, a velocità più basse, provochino la perdita di una quantità minore di energia, il processo di collisione e di coalescenza continua. Si propone l'ipotesi che il risultato possa essere una variazione radiale di opacità dell'anello.

Altre forze negli anelli

Le particelle piccole, negli anelli, possono essere soggette a forze diverse dalla gravità. Le forze elettromagnetiche sono un importante esempio. Gli anelli di Giove, di Saturno e, forse, di Urano sono immersi in un plasma a bassa densità, cioè in un gas rarefatto composto di elettroni carichi negativamente e di ioni carichi positivamente. Gli elettroni posseggono una massa inferiore a quella degli ioni, quindi si muovono più rapidamente e, inizialmente, collidono con le particelle dell'anello più frequentemente degli ioni. Accade così che le particelle assumano carica negativa in seguito all'assorbimento di elettroni. A questo punto la loro carica respinge altre particelle di carica negativa. Ancor più importante è che le particelle stesse dell'anello vengano ora accelerate da una forza elettromagnetica quando attraversano il campo magnetico del pianeta. Se le particelle sono più piccole di 0,1 micrometri circa, la forza elettromagnetica è superiore all'attrazione gravitazionale del pianeta e quindi determina il loro moto.

Molti aspetti della struttura dei sistemi di anelli potrebbero così essere spiegati. Nel caso di Giove, per esempio, l'asse del campo magnetico è inclinato di circa dieci gradi rispetto all'asse di rotazione del pianeta. In tali condizioni le forze elettromagnetiche possono impartire alle particelle piccole una distribuzione verticale molto maggiore di quella delle particelle più grandi. L'estensione verticale dell'alone del sistema di anelli di Giove è paragonabile all'estensione che ci si potrebbe aspettare per particelle di dimensioni di 0,1 micrometri o più piccole ancora. Nel caso di Saturno le forze elettromagnetiche possono rivelarsi importanti per la peculiare struttura dell'anello *F*. Inoltre, molte teorie sono state avanzate per spiegare come le razze nell'anello *B* di Saturno possano formarsi senza contemporaneamente disturbare la struttura radiale finemente suddivisa in bande dell'anello. Alcune di queste teorie fanno appello a una pioggia di particelle cariche, proveniente o dal pianeta, o dagli anelli

La presenza di anelli intorno a Urano fu desunta dagli oscuramenti registrati in occasione dell'occultazione di una stella da parte del pianeta nel 1977. La sequenza degli oscuramenti verificatisi nei minuti precedenti la scomparsa della stella dietro il pianeta *(in alto)* coincideva con quella degli oscuramenti verificatisi nei minuti seguenti la sua comparsa *(in basso)*. Per quattro oscuramenti la coincidenza era quasi esatta: quindi quattro anelli sono pressoché circolari. Un quinto, denominato epsilon, è invece molto eccentrico. Ulteriori osservazioni hanno rivelato altri quattro anelli. I dati sono stati raccolti da James L. Elliot e collaboratori con il Kuiper Airborne Observatory.

stessi, come veicolo di trasferimento di carica elettrica a particelle di dimensioni di un micrometro, o come causa di distacco delle stesse da particelle più grandi. Le particelle così staccate sono di nuovo catturate quando entrano successivamente in collisione con altre particelle grandi.

Una seconda forza che può determinare il moto di particelle piccole negli anelli è la resistenza gassosa. In questo caso l'attrito dovuto alla presenza del plasma obbliga le particelle dell'anello a scendere a spirale verso il pianeta. Più piccola è la particella, più veloce è l'abbassamento della sua orbita. In soli 20 anni circa, per esempio, la resistenza gassosa può causare l'abbassamento di una particella delle dimensioni di un micrometro dal bordo esterno a quello interno dell'anello luminoso di Giove. Altri 200 anni le occorrerebbero per scendere attraverso il disco diffuso fino all'atmosfera del pianeta. L'aspetto relativamente non strutturato dell'anello luminoso di Giove può nascere in parte da questo rapido moto radiale delle particelle piccole. In realtà, nel disco diffuso possono trovarsi particelle che provengono dall'anello luminoso.

Oltre alla gravità, alle forze elettromagnetiche e alle forze dovute alla resistenza gassosa, le particelle piccole di un sistema di anelli sono soggette a collisioni che le distruggono. Innanzitutto lo spazio interplanetario contiene corpi solidi le cui dimensioni vanno da meno di un micrometro a più di un chilometro. I corpi più piccoli sono di gran lunga i più numerosi. I corpi interplanetari sono in orbita intorno al Sole, mentre il materiale degli anelli orbita intorno a un pianeta. Perciò i due corpi possono collidere a velocità di parecchie decine di chilometri al secondo. Se il corpo interplanetario è più grande di un centesimo circa della dimensione della particella dell'anello, la collisione distrugge la particella. Su questa base si può calcolare che una particella di dimensione di un micrometro in ciascuno dei sistemi di anelli può durare solo 10 000 anni circa. Un altro tipo di collisione può essere ancora più distruttiva. Nell'anello di Giove e negli anelli esterni di Saturno è verosimile che una particella inferiore ai 10 micrometri sia erosa dalle collisioni con ioni di alta energia nella fascia di Van Allen del pianeta, ben prima che un micrometeoroide la colpisca.

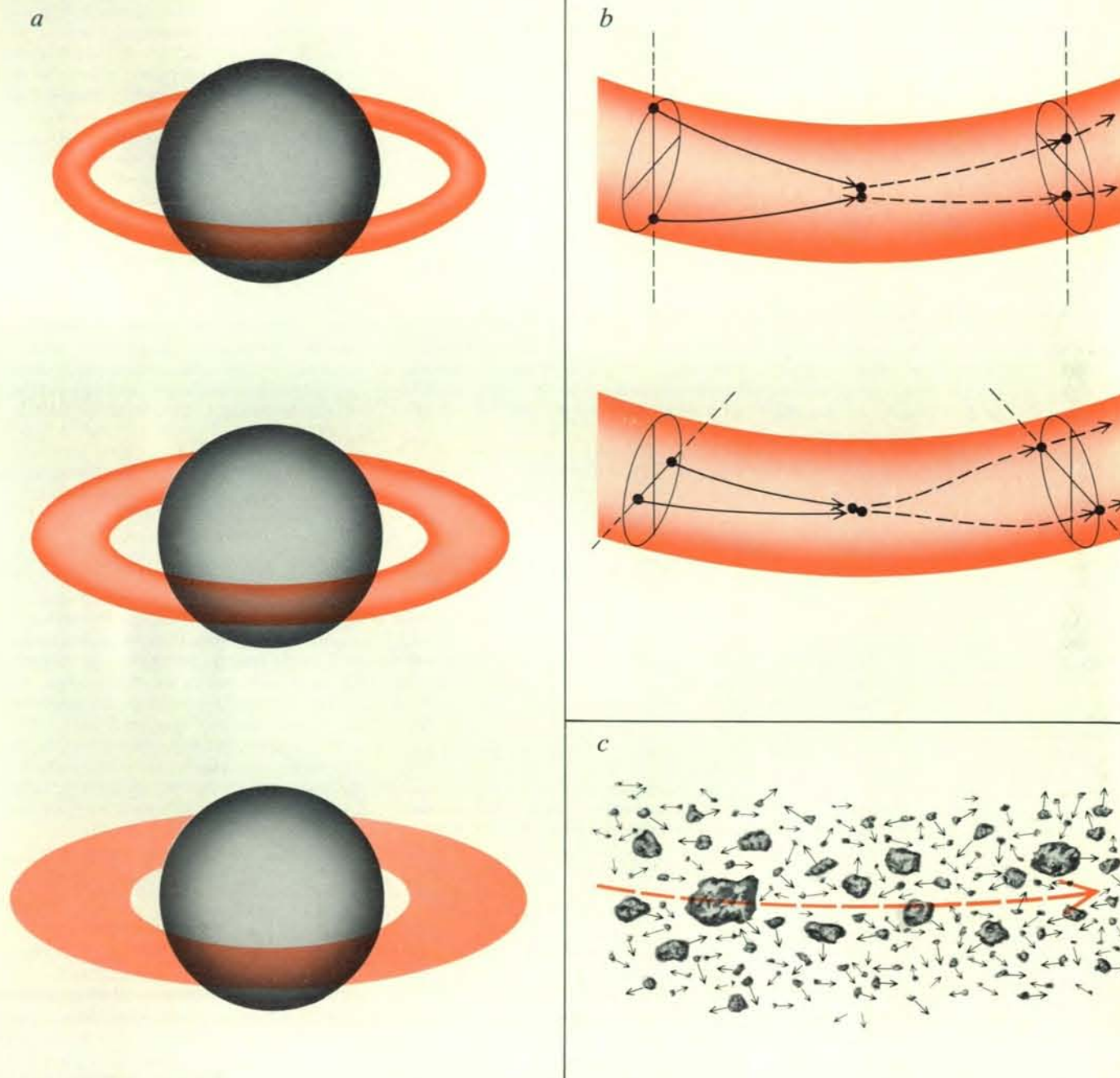
Come si sono formati gli anelli? Si ritiene che il Sole, i pianeti, i satelliti, gli asteroidi e le comete del sistema solare si siano formati circa 4,6 miliardi di anni fa all'interno della nebulosa solare, una diffusa nube di gas e polvere. Il Sole e i pianeti giganti, Giove, Saturno, Urano e Nettuno si formarono in gran parte (o forse del tutto) dai gas della nebulosa; Mercurio, Venere, Terra, Marte, Plutone, i satelliti, gli asteroidi e le comete ebbero invece origine da materia solidificata. La composizione della materia solidificata sarebbe dipesa dalla temperatura locale della nebulosa. In generale, comunque, la materia solidificata consisteva di varie mescolanze di «rocce» (inclusi silicati e ferro) e «ghiacci» (inclusa l'acqua).

Condensazione circumplanetaria

All'inizio Giove e Saturno, e forse Urano e Nettuno, erano parecchie centinaia di volte più grandi di quanto lo siano ora. Sotto l'influenza della loro stessa gravitazione gradualmente si contrassero. Man mano che ciò avveniva la velocità di rotazione aumentava. Infine la forza centrifuga, crescente in funzione della velocità di rotazione, può essere stata la causa per la quale la parte più esterna di ogni pianeta gigante in formazione - in sostanza un involucro gassoso - si separò dalla concentrazione pressoché sferica di gas al suo interno, che continuò a contrarsi. L'involucro servì come sorgente di granelli solidi dai quali ebbero origine i grandi satelliti del pianeta. Alcuni granuli si sarebbero formati per condensazione di gas, come il vapor d'acqua, a mano a mano che l'involucro si raffreddava. Forse lo stesso processo di accrescimento dei granuli in corpi più grandi diede origine a un sistema di anelli più vicino al pianeta.

Dopo un intervallo di tempo dell'ordine di un milione di anni gli involucri circumplanetari scomparvero. Essi furono forse spazzati via da un vento di gas ionizzati proveniente dal giovane Sole o forse l'attrito fra gli involucri provocò la loro caduta sui pianeti stessi. In ogni caso la formazione dei satelliti, dei satellitini e delle particelle che formano gli anelli ebbe fine.

Oggi Giove e Saturno emettono nello spazio, sotto forma di radiazioni infrarosse, all'incirca il doppio dell'energia che essi assorbono sotto forma di luce solare. La parte eccedente deriva dalla trasformazione dell'energia gravitazionale in calore per ogni contrazione del pianeta, passata e presente. All'epoca della formazione dei satelliti, i pianeti si stavano contraendo molto più rapidamente. Per questo Giove e Saturno e forse gli altri pianeti giganti possono aver irradiato calore in quantità molte migliaia di volte superiore a quanto ne emettano ora. Questo calore può ben aver determinato la temperatura e quindi la composizione della materia solida negli involucri circumplanetari.

Molte ipotesi sull'origine dei sistemi di anelli intorno a tre pianeti giganti gassosi sostengono che il materiale dell'anello si sia formato negli involucri circumplanetari in posizioni prossime alla sua attuale distanza radiale. Variazioni di temperatura e delle altre condizioni nei pressi di ciascun pianeta possono essere state la causa delle notevoli differenze nella composizione dei satelliti vicini e del materiale dei loro anelli. Le densità medie dei quattro satelliti maggiori di Giove indicano che i due satelliti interni sono costituiti quasi esclusivamente di rocce e gli altri

L'evoluzione dei sistemi di anelli ebbe probabilmente inizio nel sistema solare primitivo, quando una nube gassosa in rotazione, contraendosi in un pianeta gigante, come Saturno, lasciò alle proprie spalle un involucro gassoso spinto all'esterno dalla forza centrifuga dovuta alla rapida rotazione *(a)*. Le particelle in accrescimento nell'involucro dovevano essere animate da moto orbitale intorno al pianeta, ma anche da movimenti casuali che provocavano collisioni *(b)*. Le collisioni erano anelastiche e determinavano nelle particelle deformazioni e riscaldamento. La conseguente perdita di energia cinetica diminuiva il loro moto verticale. Le collisioni fra due particelle a distanza radiale prossima provocavano un ulteriore effetto. In ciascuna collisione di quel tipo la particella interna percorreva la propria orbita a velocità maggiore dell'esterna e colpiva questa da dietro. Il risultante trasferimento di quantità di moto aveva per effetto di innalzare l'orbita della particella esterna e di abbassare l'orbita dell'interna. Perciò l'involucro si espandeva orizzontalmente. Ne risultò un disco piatto nel quale le velocità casuali delle particelle, l'una rispetto all'altra, non superavano un decimilionesimo del valore delle loro velocità orbitali medie *(c)*. Si ritiene che le particelle più piccole che oggi si trovano negli anelli siano il frutto di erosione derivante da collisioni di particelle di dimensioni maggiori.

due di equivalenti quantità di rocce e ghiaccio. Poiché l'anello di Giove è ben all'interno dell'orbita del grande satellite più vicino al pianeta, qualunque particella che si fosse formata presso l'anello non poteva essere costituita che di roccia; anzi, poteva essere costituita solo di minerali relativamente rari che solidificano ad alte temperature.

Al contrario, i grandi satelliti più interni di Saturno sono composti tutti essenzialmente di ghiaccio. Quindi può darsi che le temperature, anche a breve distanza da Saturno, dove ora giacciono gli anelli del pianeta, fossero abbastanza basse da permettere la formazione di particelle composte, in grande parte, di ghiaccio. Le differenze fra le condizioni intorno a Giove e quelle intorno a Saturno sono presto spiegate. Prima di tutto, la quantità di calore generata nei tempi primitivi dalla contrazione di una sfera gassosa dipende approssimativamente dal quadrato della sua massa. Giove, con una massa di oltre tre volte quella di Saturno, generò dieci volte più calore. Può anche darsi che presso Saturno non fosse disponibile tanto materiale roccioso da costituire particelle. Al momento della formazione dell'involucro circumplanetario di Saturno, l'atmosfera del pianeta si estendeva probabilmente fino alla regione ora occupata dagli anelli e dai satelliti più interni. Da ciò consegue che la maggior parte dei granuli di roccia possono essere stati trattenuti dall'atmosfera di Saturno, oppure possono essere stati richiamati all'esterno dai satelliti.

Le ipotesi che presuppongono basse temperature e scarsità di rocce presso Saturno trovano avallo nel fatto che i satelliti più interni di Saturno hanno una densità media più bassa e conseguentemente un maggiore contenuto di ghiaccio della maggior parte dei satelliti più esterni. Tali ipotesi ci aiutano anche a spiegare perché gli anelli di Saturno sembrino essere formati quasi esclusivamente di ghiaccio, mentre quelli di Giove siano rocciosi. Sembra che anche gli anelli di Urano siano di roccia, mentre si sarebbe potuto pensare che fossero di ghiaccio.

Ipotesi rivali

Le varie ipotesi che danno una spiegazione degli anelli intorno ai pianeti giganti gassosi differiscono in primo luogo per l'importanza attribuita alla relazione fra le particelle oggi osservate e il primitivo materiale dell'anello. Secondo un'ipotesi, un singolo grosso corpo si sarebbe spezzato in una miriade di pezzi mentre si avvicinava al pianeta e i frammenti avrebbero formato gli anelli. Può darsi che il corpo fosse un grosso meteoroide, soggetto a un casuale incontro gravitazionale col pianeta, o poteva anche essere un satellitino formatosi nell'involucro del pianeta. In entrambi i casi la frammentazione sarebbe stata dovuta all'effetto di marea: la forza di taglio che ha origine per l'attrazione gravitazionale esercitata dal pianeta sul corpo è maggiore per le parti del corpo vicine al pianeta e inferiore per le più lontane. La creazione degli anelli di Saturno per l'effetto di marea fu proposta dal matematico francese Edouard Albert Roche nel 1848. Egli aveva calcolato che le forze di marea avrebbero superato quelle di coesione per autogravitazione di un satellite liquido nel caso si fosse avvicinato a meno di un raggio e mezzo di Saturno dalla superficie del pianeta. Questa soglia di frammentazione - il limite di Roche - si trova presso il bordo esterno degli anelli principali di Saturno.

È del tutto inverosimile, però, che un satellitino vicino a Saturno fosse liquido. Avrebbe dovuto essere solido e un satellitino solido è tenuto insieme non solo dalla autogravitazione ma anche dalle forze che legano gli atomi in cristalli. Secondo Hans R. Aggarwal e Verne Oberbeck, dell'Ames Research Center della NASA, un satellitino solido di diametro inferiore ai 100 chilometri non può essere distrutto per effetto di marea a qualunque distanza dalla superficie del pianeta. Inoltre, un satellite solido di dimensioni superiori non può essere frammentato a una distanza superiore ai quattro decimi del raggio del pianeta dalla superficie. Tale distanza pone la soglia di frammentazione entro il bordo interno degli anelli principali di Saturno. È anche inverosimile che l'effetto di marea su un meteoroide vagante possa aver prodotto le particelle dell'anello vicino a Saturno. Queste, come il corpo generatore, avrebbero avuto velocità sufficienti per sfuggire dalle vicinanze del pianeta.

C'è però un modo sottile nel quale le forze di marea avrebbero potuto essere importanti. Come ha dimostrato Roman Smoluchowski, dell'Università del Texas ad Austin, l'attrazione gravitazionale di particelle di uguale dimensione è insufficiente a tenerle unite contro la frammentazione dovuta a forze di marea all'interno del classico limite di Roche (il limite calcolato da Roche per un satellite liquido). Al contrario, due particelle che differiscono molto in dimensione possono resistere alla frammentazione da marea a distanze ben inferiori al limite classico. Nella realtà, il limite definito per la frammentazione da marea di due particelle di uguale dimensione e quello definito per due particelle di dimensioni diverse si trovano presso i bordi esterno e interno dei sistemi di anelli intorno a Giove, Saturno e Urano. Entro il limite esterno di frammentazione le particelle possono aver subito un processo di accrescimento, ma così lentamente da non diventare mai grossi satelliti. Entro il limite di frammentazione interno l'accrescimento può essere stato quasi impossibile.

La seconda tra le ipotesi più valide a proposito della storia delle particelle di un anello fu esposta da Eugene Shoemaker, dell'US Geological Survey. Essa postula che un unico grande satellite nella regione degli anelli (o forse un certo numero di satelliti) abbia subito una collisione catastrofica con un meteoroide vagante. Le fotografie dei satelliti di Giove e di Saturno riprese dai Voyager mettono infatti in rilievo su quei satelliti un gran numero di crateri prodotti da collisioni ad alta velocità. Mimas, uno dei satelliti più piccoli di Saturno, ha un cratere che occupa un terzo di emisfero. La collisione che produsse questo cratere deve essere stata tale da rasentarne la distruzione. Più vicini a Saturno, due corpi con un diametro di circa 100 chilometri ruotano in orbite pressoché identiche. Potrebbero essere i più grossi frammenti prodotti da una collisione catastrofica fra un satellite e un meteoroide gigante.

Vi sono molte ragioni per pensare che collisioni catastrofiche abbiano avuto occasione di verificarsi preferenzialmente nella regione che avrebbe dovuto, in seguito, essere occupata dagli anelli. Innanzitutto i principali satelliti di Giove e di Saturno tendono ad avere dimensioni minori a mano a mano che ci si avvicina al pianeta. In corrispondenza di un certo livello di energia di collisione è più probabile che si frammentino i piccoli satelliti anziché i grandi. Inoltre il campo gravitazionale di un pianeta determina le traiettorie dei meteoroidi in modo tale da incrementare il flusso di meteoroidi in prossimità del pianeta rispetto al flusso degli stessi a distanze crescenti.

Un'ultima ipotesi riguardante la storia delle particelle degli anelli postula che i corpi più grossi negli anelli siano semplicemente il prodotto dello scarso accrescimento di materia nell'involucro circumplanetario a distanze prossime al pianeta. L'accrescimento ebbe inizio col raffreddamento dell'involucro e con la conseguente condensazione di massa gassosa in minuscoli granuli solidi. Le forze gravitazionali e la resistenza gassosa obbligarono i granuli a disporsi sul piano equatoriale dell'involucro. Ivi i i granuli continuarono a crescere per condensazione di vapore sulla loro superficie. Tale crescita permise loro di raggiungere dimensioni di alcuni metri. Le particelle che costituiscono la maggior parte degli anelli visibili di Saturno hanno dimensioni che variano da alcuni centimetri ad alcuni metri e, quindi, possono rappresentare il risultato di tale processo. Ogni satellitino negli anelli sarebbe quindi il prodotto locale di uno stadio di crescita in cui corpi di dimensione dell'ordine del metro si accrebbero ulteriormente in seguito a collisione morbida.

Le particelle in tutti i tre sistemi di anelli sono piccole e numerose e l'ipotesi dell'accrescimento trova valide giustificazioni. Innanzitutto la formazione indisturbata di minuscoli granuli non avrebbe potuto incominciare, a una data distanza da un pianeta gigante in formazione, finché il pianeta non si fosse ridotto di dimensione entro quella data distanza e l'involucro circumplanetario non si fosse sufficientemente raffreddato. Meno tempo era quindi disponibile per la formazione di granuli vicino al pianeta che non a distanze maggiori. Inoltre i minuscoli granuli vicini a Giove potevano formarsi soltanto dai relativamente rari prodotti disponibili di condensazione con punto di solidificazione ad alta temperatura. Infine, il pervenire di un piccolo satellite a una certa dimensione significherebbe che le sue risonanze sovrapposte avevano una ampiezza paragonabile alle sue dimensioni. Nuovo materiale non era più in grado di aggiungersi, perciò la crescita cessava. Negli anelli di Saturno il limite di crescita calcolato va da alcuni chilometri fino a qualche decina di chilometri. Condizioni diverse devono essersi avute a distanze maggiori da Saturno e dagli altri pianeti giganti, dal momento che là si formarono satelliti di maggiori dimensioni.

Erosione continua

Che cosa si può dire a proposito delle piccole particelle nei sistemi di anelli? Abbiamo già notato che la resistenza gassosa obbliga le particelle di dimensioni micrometriche a percorrere una traiettoria a spirale dal bordo esterno dell'anello luminoso fino all'atmosfera di Giove in un tempo di poche centinaia di anni. È chiaro che tali particelle non hanno potuto sopravvivere dal tempo in cui il pianeta aveva un involucro gassoso. Esse quindi devono formarsi ancora oggi. Secondo Joseph A. Burns, della Cornell University, esse sono il prodotto dell'erosione di corpi più grossi all'interno o vicino all'anello luminoso di Giove.

Le particelle degli anelli più grandi di un centimetro, poco più, poco meno, non vengono facilmente distrutte in seguito a collisioni con micrometeoroidi interplanetari, ma ogni collisione provoca un minuscolo cratere intorno al punto di impatto e una quantità di materia con una massa da 1000 a 10 000 volte più grande di quella del corpo che entrò in collisione viene espulsa. Molte particelle di dimensioni micrometriche presenti nei sistemi di anelli possono essere il frutto di queste espulsioni. Si può stimare che se un satellitino ha un diametro inferiore a 10 chilometri circa, la maggior parte di queste particelle, espulse in seguito a collisione con un corpo interplanetario, può sfuggire al campo gravitazionale del piccolo satellite. Le particelle espulse, che sfuggono al satellitino, mancherebbero dell'energia necessaria per sfuggire al pianeta intorno al quale il satellite orbita e si stabilizzerebbero quindi negli anelli. (La continua erosione dei corpi originari potrebbe essere anche l'origine, negli anelli di Saturno, di corpi di dimensioni centimetriche o anche di un metro.)

In definitiva, sia i satellitini sia le particelle più grandi degli anelli hanno avuto origine fin dai primordi del sistema solare: sono contemporanei dei satelliti dei pianeti giganti. Le particelle più piccole degli anelli si stanno formando ancora oggi. Si suppone che i grandi satelliti del sistema solare esterno, come pure alcuni pianeti (inclusa la Terra), si siano formati in seguito a un processo di accumulazione di molti corpi di dimensioni minori. Sicuramente, fra le molte particelle degli anelli, tale processo su piccola scala è tutt'ora in atto. I sistemi di anelli pongono quindi una doppia sfida: da un lato si cerca di conoscere i processi che portarono alla loro formazione, dall'altra si cerca di servirsi di queste conoscenze per chiarire la formazione dei satelliti e dei pianeti solidi.

La riparazione inducibile del DNA

Un danno al materiale ereditario attiva una «risposta di tipo SOS», in cui parecchi sistemi enzimatici funzionano assieme. Due proteine, RecA e LexA, svolgono molti ruoli nella riparazione di questo danno genetico

di Paul Howard-Flanders

La vita di ogni organismo e la sua continuità di generazione in generazione dipendono dalla stabilità a lungo termine dell'informazione ereditaria, codificata nella doppia elica di DNA. Questa stabilità non è il risultato di una invulnerabilità al danno: il DNA, sia che si trovi in una cellula batterica sia che appartenga ai milioni di cellule che costituiscono un essere umano, è sensibile al danno provocato dalle radiazioni e dagli agenti chimici che sono presenti nell'ambiente. La sua stabilità e la precisione con la quale si replica sono mantenute da enzimi che, ininterrottamente, riparano le lesioni genetiche. In particolare, un danno significativo alle molecole di DNA può indurre una risposta di emergenza in cui vengono sintetizzate maggiori quantità degli enzimi di riparazione, facendole funzionare in un insieme ben concertato. Ormai da molti anni si osservano e si studiano vari sistemi di riparazione del DNA, ma solo negli ultimi due anni si è cominciato a comprendere la natura di certe risposte indotte e di importanti sistemi di regolazione che le controllano.

Sia la vulnerabilità al danno del DNA, sia la sua suscettibilità alla riparazione sono implicite nella struttura stessa del DNA. Ogni filamento della doppia elica è una catena di molecole di zucchero e di gruppi fosfato. Da questa impalcatura di sostegno sporge una serie di composti chimici, chiamati basi. Le effettive unità costruttive del DNA sono però i nucleotidi, ciascuno costituito da una base, da uno zucchero e da un gruppo fosfato. Le basi sono di quattro tipi: due purine, l'adenina *(A)* e la guanina *(G)*, e due pirimidine, la citosina *(C)* e la timina *(T)*. Le basi presenti in un filamento sono unite da legami idrogeno alle basi presenti sull'altro filamento e formano così il DNA a doppio filamento. Le basi sono complementari: *A* si appaia solo con *T*, e *G* solo con *C*. Di conseguenza, i filamenti sono tra di loro complementari e ciascuno è, di fatto, uno stampo per l'altro: laddove un filamento ha la sequenza *TGACCGCTTA*, l'altro deve avere la sequenza *ACTGGCGAAT*. La complementarietà è la base della replicazione del DNA, della sua espressione in forma di proteina e della sua ricombinazione. Nella replicazione, ogni filamento serve da stampo sul quale viene deposto un nuovo filamento. Nell'espressione del DNA, l'informazione ereditaria, codificata nella sequenza di basi su un filamento di un gene, viene trascritta con la sequenza complementare presente su un filamento di RNA e quindi tradotta in proteina; ogni «codone» di tre basi specifica uno degli amminoacidi che costituiscono la catena proteica. Nella ricombinazione genetica, alcuni filamenti di DNA vengono tagliati mentre regioni omologhe si uniscono producendo nuove combinazioni di geni.

Evidentemente l'incorporazione nel DNA di una base sbagliata o alterata, o la presenza di una lesione qualsiasi che distorca la doppia elica o impedisca un accurato appaiamento delle basi, tenderà a interferire con la replicazione, con la sintesi proteica e con la ricombinazione. A meno che non venga riparato bene, il danno al DNA può provocare la morte di una cellula; sembra, talvolta, che nelle cellule di un mammifero esso inneschi un processo che conduce al cancro.

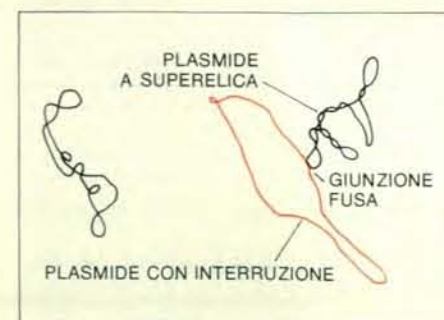
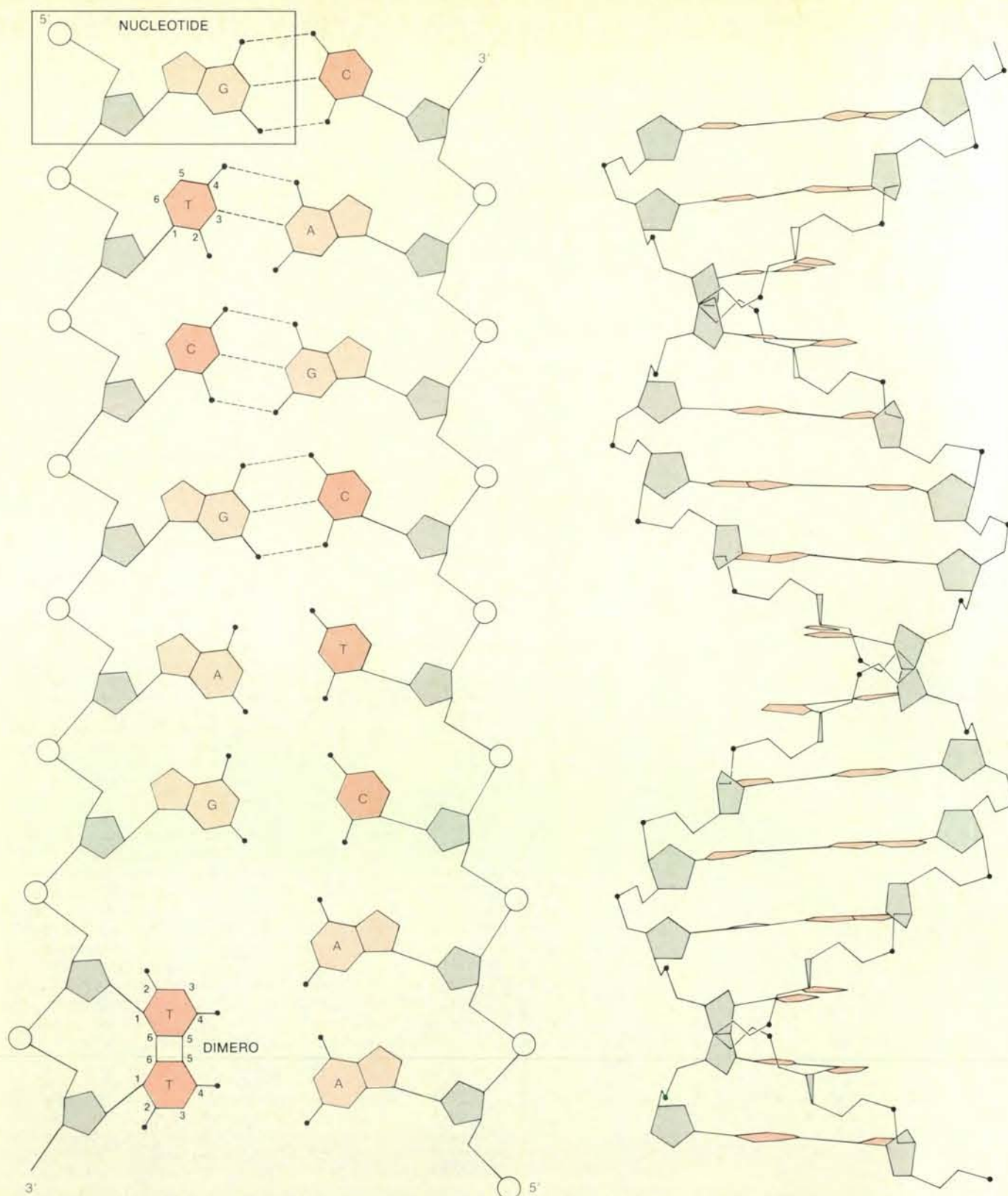
Il danno al DNA può assumere molte forme, ma la lesione la cui riparazione è stata studiata più accuratamente è il dimero pirimidinico, un prodotto dovuto all'esposizione alle radiazioni ultraviolette. Un dimero si forma quando due basi pirimidiniche adiacenti (o timine o citosine) su uno stesso filamento sono legate tra loro da un anello a quattro atomi di carbonio. Le due basi di un dimero vengono spinte fuori dall'allineamento, per cui i loro legami idrogeno con le basi complementari si rompono e l'impalcatura di sostegno del DNA si distorce, il che a sua volta impedisce il corretto appaiamento delle due basi su ciascun lato del dimero. La presenza di un unico dimero pirimidinico è sufficiente a far interrompere sia la trascrizione sia la replicazione.

Le nozioni fondamentali sui processi di riparazione del DNA sono state ricavate, in primo luogo, da esperimenti genetici con cellule batteriche esposte alle radiazioni ultraviolette e da studi biochimici sul corso seguito dalla replicazione e dalla riparazione del DNA e sugli enzimi che partecipano a questi processi. In questo articolo, tracceró cronologicamente la storia di alcune scoperte, nella speranza di riuscire a ricordare al lettore che i sistemi biologici complessi non sono spesso compresi tutti insieme o addirittura secondo una progressione logica. Capita piuttosto, di solito, che si accumulino, spesso senza ovvie relazioni o immediato significato, numerose osservazioni e frammenti di dati, che acquistano significato l'uno dall'altro: con il procedere delle ricerche i frammenti si riuniscono e cominciano ad assumere un senso.

Già negli anni cinquanta si era osservato che, mentre il comune batterio intestinale *Escherichia coli* sopravvive in genere a moderate esposizioni alle radiazioni ultraviolette, un ceppo mutante del batterio risulta particolarmente sensibile, o vulnerabile, a tale radiazione. Perché? Una possibile spiegazione deriva dal fatto che i batteri normali dispongono di un enzima, o di un sistema di enzimi, in grado di riparare i danni al DNA, provocati dai raggi ultravioletti, aumentando così il tasso di sopravvivenza, mentre i batteri sensibili sarebbero privi di questa capacità. Gli enzimi, come altre proteine, sono codificati da geni e, nel 1961, i miei colleghi e io avviammo degli esperimenti di genetica alla Yale University School of Medicine per trovare geni che potessero specificare gli enzimi di riparazione. Fummo avvantaggiati dal fatto che le cellule di *E.coli* possono scambiarsi geni per coniugazione, una forma primitiva di accoppiamento tra cellule maschili e femminili. Abbiamo lavorato con il ceppo standard di laboratorio *E.coli* K12, perché era nota la sua relativa resistenza alle radiazioni ed era stata determinata la sua mappa genetica, cioè la localizzazione di molti geni lungo la sua molecola circolare a duplice filamento di DNA.

Nell'isolare i mutanti di *E.coli* sensibili alle radiazioni ultraviolette, ai raggi X o ad ambedue, abbiamo trovato che la loro sensibilità era una proprietà ereditaria, stabile, che si poteva trasmettere tra cellule in accoppiamento. Con metodi genetici standard siamo stati in grado di identificare i geni mutanti e di localizzare le loro posizioni sul cromosoma di *E.coli*. Abbiamo dimostrato che una classe di mutanti sensibili alle radiazioni ultraviolette possedeva mutazioni nei geni che designammo con *uvrA*, *uvrB* e *uvrC*; altri mutanti, sensibili sia alle radiazioni ultraviolette sia ai raggi X, presentavano mutazioni in un gene che chiamammo *lexA*. I geni *uvr* - come risultò evidente - codificano per gli enzimi di riparazione; il gene *lexA* per un differente tipo di proteina, ma questo non risultò chiaro che molto tempo dopo.

Quale meccanismo biochimico è mediato dai geni *uvr?* Richard B. Setlow e William L. Carrier dell'Oak Ridge National Laboratory e P. Boyce, Lee Theriot e io stesso della Yale University abbiamo studiato il destino dei dimeri pirimidinici nel DNA di *E.coli* normale e di mutanti sensibili alle radiazioni ultraviolette, *uvr* negativi, in quanto carenti dell'uno e dell'altro gene *uvr*. Per far questo, abbiamo esposto le cellule del batterio alle radiazioni ultraviolette e quindi abbiamo esaminato il loro DNA alla ricerca dei dimeri pirimidinici. Abbiamo trovato che, mentre essi permanevano nel DNA di cromosomi intatti di cellule *uvr* negative (come ci aspettavamo), non erano mai presenti, invece, nel DNA cromosomico dei batteri resistenti. I dimeri non erano però scomparsi da tali cellule: si potevano, infatti, trovare in brevi filamenti singoli di DNA, costituiti solo da alcune basi. Ciò che avevamo scoperto era la «riparazione per excisione» (ossia per eliminazione). I singoli filamenti di DNA che incorporavano i dimeri erano stati excisi, presumibilmente da un enzima o da enzimi codificati da uno o più geni *uvr*. Abbiamo pensato allora che la riparazione del DNA potesse essere portata a termine dagli enzimi DNA-polimerasi e DNA-ligasi, che rispettivamente sostituirebbero i nucleotidi eliminati e si legherebbero all'impalcatura zucchero-fosfato. Il filamento complementare integro fungerebbe da stampo. Ben presto dimostrazioni dirette di un processo di riparazione di questo tipo furono riportate da David E. Pettijohn e da Philip C. Hanawalt della Stanford University.

È stato trovato che la riparazione per excisione è uno dei processi fondamentali di riparazione del DNA non solo nei batteri, ma anche nelle cellule dei mammiferi (comprese quelle umane), ed è in grado di

La ricombinazione, mediata dalla proteina RecA, è la base di un meccanismo di riparazione del DNA. Una fase del processo di ricombinazione viene illustrata nella microfotografia elettronica scattata da Era Cassuto e da Stephen C. West nel laboratorio dell'autore. Come viene indicato nella mappa della microfotografia *(a sinistra)*, due copie di un plasmide di DNA (una piccola molecola circolare a doppio filamento) vengono mostrate mentre si appaiano in corrispondenza di una «giunzione fusa», un evento precoce della ricombinazione. Un plasmide è una molecola circolare, integra, di DNA a duplice filamento, come risulta dalla sua forma avvolta a «superelica». È appaiato con un altro plasmide che è identico, tranne per il fatto che un segmento breve di un filamento del suo DNA è stato rimosso da enzimi, lasciando una interruzione (che fa «decontrarre» il plasmide). Queste coppie fuse vennero trovate solo quando plasmidi integri e altri analoghi con una interruzione a filamento singolo furono incubati in presenza della proteina RecA e di ATP; le coppie includevano sempre almeno una molecola decontratta. Presumibilmente la proteina RecA si legava alla regione a filamento singolo, opposta alla interruzione, e la allineava con una regione omologa, presente in un filamento del plasmide intatto.

La molecola di DNA *(a sinistra)* consiste di due filamenti di nucleotidi, ciascuno comprendente uno zucchero *(in grigio)*, un gruppo fosfato *(cerchio vuoto)* e una delle quattro basi *(in colore)*: le purine adenina *(A)* e guanina *(G)* e le pirimidine timina *(T)* e citosina *(C)*. Le basi sporgono dall'impalcatura zucchero-fosfato e i due filamenti sono uniti da legami idrogeno *(linee tratteggiate)* tra basi complementari: *A* si appaia solo con *T*, e *G* con *C*. L'informazione genetica è codificata nella sequenza delle basi di un filamento. I due filamenti hanno polarità opposta; la sequenza di basi si legge per convenzione nella direzione 5'-3'. In questo schema la doppia elica appare srotolata; il piano delle basi, che è effettivamente perpendicolare rispetto all'asse dell'elica, è ruotato di 90 gradi in modo da mostrare la struttura chimica. I dimeri pirimidinici, cioè le lesioni che vengono maggiormente studiate nelle indagini sulla riparazione del DNA, sono provocati dall'esposizione alle radiazioni ultraviolette. Un dimero *(in basso)* si forma quando i doppi legami tra gli atomi di carbonio 5' e 6' di due pirimidine adiacenti si aprono e si forma un anello a quattro atomi di carbonio, con le due basi che vengono trascinate insieme e i loro legami idrogeno con le basi complementari sul filamento opposto che si rompono; l'impalcatura della molecola si distorce e non è più possibile una corretta formazione di legami idrogeno nemmeno per due basi che si trovino su ogni lato del dimero. I dimeri si formano il più delle volte tra timine adiacenti su uno stesso filamento, ma anche tra due residui di citosina o tra una timina e una citosina. Un modello di impalcatura di DNA *(a destra)* mostra come le due catene zucchero-fosfato sono avvolte a doppia elica: qui le basi sono quasi perpendicolari al piano della pagina.

eliminare non solo i dimeri pirimidinici, ma anche altri tipi di lesioni al DNA. Per esempio, Boyce e io abbiamo trovato con facilità che essa ripara il danno provocato dalla mitomicina *C*, che agisce in primo luogo formando legami trasversali tra i due filamenti della doppia elica. Il fatto che la mitomicina *C* sia un farmaco antitumorale (in quanto danneggia il DNA) e sia anche (per la stessa ragione) cancerogeno nelle cellule animali servì a ricordare che le nozioni acquisite dagli studi sulla riparazione del DNA nei batteri potevano avere applicazioni molto ampie (si veda anche in «Le Scienze» l'articolo *Geni discontinui* di Pierre Chambon, n. 155, luglio 1981).

Alla scoperta di un secondo importante processo di riparazione del DNA, molto diverso dalla riparazione per excisione, si giunse per via indiretta. Nel 1965, A. John Clark e Anne D. Margulies della Università della California a Berkeley stavano studiando la ricombinazione genetica, il processo in cui geni situati su cromosomi diversi si associano, ad esempio nel corso dell'accoppiamento batterico. Clark isolò mutanti di *E.coli* il cui DNA non riusciva a ricombinarsi come ci si aspettava, dopo l'accoppiamento, e fece risalire questa incapacità di ricombinarsi a una mutazione in un gene particolare, che chiamò *recA*. Quando sottopose le cellule *recA* negative a una prova per la resistenza alle radiazioni, trovò con sua sorpresa che esse erano estremamente sensibili al danno da radiazioni ultraviolette e da raggi X.

W. Dean Rupp e io stesso fummo imbarazzati di fronte a questa evidente relazione tra assenza di ricombinazione e sensibilità alle radiazioni. Forse ciò voleva dire che il prodotto proteico di *recA* mediava qualche processo di riparazione fino ad allora sconosciuto. Confrontammo la sensibilità all'ultravioletto di cellule prive di *uvrA* o di *recA* con la sensibilità dei doppi mutanti, mancanti di ambedue i geni suddetti. I doppi mutanti erano una cinquantina di volte più sensibili, o delle cellule *uvrA* negative, o delle cellule *recA* negative, una chiara indicazione questa che vi erano interessati due differenti processi di riparazione. Per semplificare le cose, studiammo la sintesi del DNA nei batteri *uvrA* negativi, che non potevano effettuare la riparazione per excisione, ma presumibilmente conservavano qualunque sistema fosse basato sul prodotto del gene *recA*. Nel 1968 scoprimmo la riparazione per postreplicazione, che si realizza solo dopo la replicazione del DNA danneggiato dalle radiazioni ultraviolette.

In una cellula esposta alle radiazioni ultraviolette, il DNA si replica normalmente nei tratti non lesi dei filamenti progenitori che fungono da stampo. Quando la biforcazione per la replicazione (cioè il punto in cui i due filamenti progenitori si srotolano e rimangono esposti agli enzimi della replicazione) raggiunge un dimero pirimidinico che non è stato exciso, non è possibile lungo il filamento danneggiato un corretto appaiamento di basi. La sintesi di un nuovo filamento viene così interrotta, per ricominciare in qualche punto del tratto non leso dello stampo, oltre il dimero. Come conseguenza, si formano nei filamenti figli delle «interruzioni da postreplicazione», le quali non possono essere sottoposte a riparazione per excisione, che richiede un filamento complementare intatto: devono, così, essere riparate con un processo diverso. Abbiamo pensato che la sequenza delle basi andasse effettivamente perduta in ambedue i filamenti di uno dei due «duplex» (cioè doppio filamento che si forma dalla ricombinazione) appena formati: un filamento aveva una interruzione e l'altro un dimero. L'informazione mancante doveva essere in qualche modo recuperata perché la interruzione potesse essere riparata correttamente. Ci siamo accorti che la sequenza di basi necessaria era disponibile lì vicino, nel duplex fratello formato in corrispondenza della stessa biforcazione per la replicazione. Poteva, dunque, un evento di ricombinazione sostituire, in uno dei due filamenti del duplex fratello, le basi mancanti nella interruzione da postreplicazione, fornendo così uno stampo, con la giusta sequenza di basi, opposto al dimero?

Nel tentativo di scoprire questa riparazione per ricombinazione, Rupp e io marcammo il DNA con isotopi, il che permise di distinguere, gli uni dagli altri, i filamenti progenitori e quelli figli sulla base della loro densità. Fummo in grado di contare il numero delle interruzioni da postreplicazione e il numero di scambi tra filamenti fratelli. Possiamo affermare con relativa sicurezza che il numero di scambi tra filamenti fratelli e interruzioni da postreplicazione era all'incirca lo stesso. La conferma della riparazione per ricombinazione da postreplicazione venne da parte dei risultati ottenuti da Ann K. Ganesan di Stanford, in base ai quali i dimeri pirimidinici non rimangono soltanto nei filamenti trattati con radiazioni ultraviolette, ma vengono distribuiti anche tra i nuovi filamenti, così come tra i vecchi, allorché avviene la replicazione del DNA.

La riparazione da postreplicazione opera bene in complementarietà con la riparazione da excisione. Le interruzioni lasciate nei filamenti figli dai dimeri (o da altre forme di danno), che raggiungono la biforcazione per la replicazione prima di essere excisi, sono reintegrate mediante riparazione per ricombinazione; il dimero (o altro danno) presente in un filamento progenitore viene sottoposto a riparazione per excisione.

Il danno al DNA si può manifestare in molte forme. Come viene indicato sulla doppia elica *(in alto)*, una base può venire alterata chimicamente per alchilazione o ossidrilazione o in altri modi ancora *(1)* oppure può venire eliminata *(2)*. Per aggiunta di una molecola voluminosa si può formare un addotto chimico *(3)*. Le pirimidine adiacenti possono unirsi in un dimero *(4)*, i due filamenti possono incrociarsi grazie a legami chimici *(5)* oppure uno dei due filamenti può legarsi a una proteina. Un filamento *(6)* o entrambi *(7)* possono anche rompersi. La maggior parte di questi danni provoca la rottura dell'appaiamento delle basi o della continuità dello stampo e, pertanto, interferisce con il corretto funzionamento del DNA. La replicazione e la trascrizione dipendono dall'appaiamento delle basi per la sintesi di due nuovi filamenti figli di DNA o per la sintesi di un filamento di RNA messaggero; la maggior parte dei processi di ricombinazione dipende anch'essa da tale appaiamento delle basi per riuscire a far allineare tratti omologhi di DNA.

La riparazione per excisione rimuove i dimeri pirimidinici e qualche altro danno a carico delle basi dal DNA, mediante un processo di «taglia e cuci». Tre enzimi speciali, codificati dai geni *uvrA*, *uvrB* e *uvrC*, vi prendono parte assieme agli enzimi DNA-polimerasi I e DNA-ligasi. In primo luogo, le proteine sintetizzate sotto il controllo dei geni *uvrA* e *uvrB* si legano al sito danneggiato *(1)* e forse, con l'aiuto della proteina sintetizzata sotto il controllo del gene *uvrC*, producono una spaziatura *(freccia in colore)* in corrispondenza della estremità 5' della regione lesa *(2)*. In presenza della proteina sintetizzata sotto il controllo del gene *uvrC*, la DNA-polimerasi I si lega in corrispondenza della spaziatura e, in accordo con le regole di appaiamento delle basi, aggiunge nucleotidi alla estremità 3'. Essa produce nella catena una seconda spaziatura per poter procedere all'eliminazione della regione danneggiata *(3)*; può poi continuare a digerire i nucleotidi e a sostituirli uno alla volta, spostando progressivamente la spaziatura verso destra *(4)*. Infine, la spaziatura *(freccia spessa)* viene chiusa dalla DNA-ligasi, e così si completa la riparazione.

Una maggiore comprensione del ruolo svolto dal *recA* nella riparazione per postreplicazione è stata raggiunta in seguito a studi compiuti su un fenomeno che non aveva apparentemente nessuna relazione con essa: l'induzione lisogena. Si dice che un batterio è lisogeno quando il DNA di un batteriofago (virus batterico), che lo infetta, si integra nel suo cromosoma batterico come profago invece di moltiplicarsi. (Una simile moltiplicazione porterebbe alla formazione di nuove particelle fagiche che ucciderebbero la cellula.) Uno di questi virus «temperati», che infetta ceppi lisogeni di *E.coli*, è il fago lambda. Il suo DNA può rimanere quiescente nel batterio, replicandosi con il cromosoma batterico per molte generazioni. Quando il batterio si trova esposto alle radiazioni ultraviolette o a qualche altro agente che danneggia il DNA, il profago viene «indotto»: il suo DNA forma un'ansa all'esterno del cromosoma, si replica e dirige la sintesi delle proteine virali; centinaia di nuove particelle virali vengono prodotte ed emesse dalla cellula che scoppia. Nel 1967, Clark, che aveva scoperto il *recA*, e Israel Hertman con Salvador E. Luria, del Massachusetts Institute of Technology, trovarono che le cellule lisogene *recA* negative non vengono indotte, cioè il fago lambda rimane un profago quiescente anche quando la cellula viene esposta alle radiazioni ultraviolette. Il gene *recA*, in altre parole, risulta essere necessario per l'induzione lisogena.

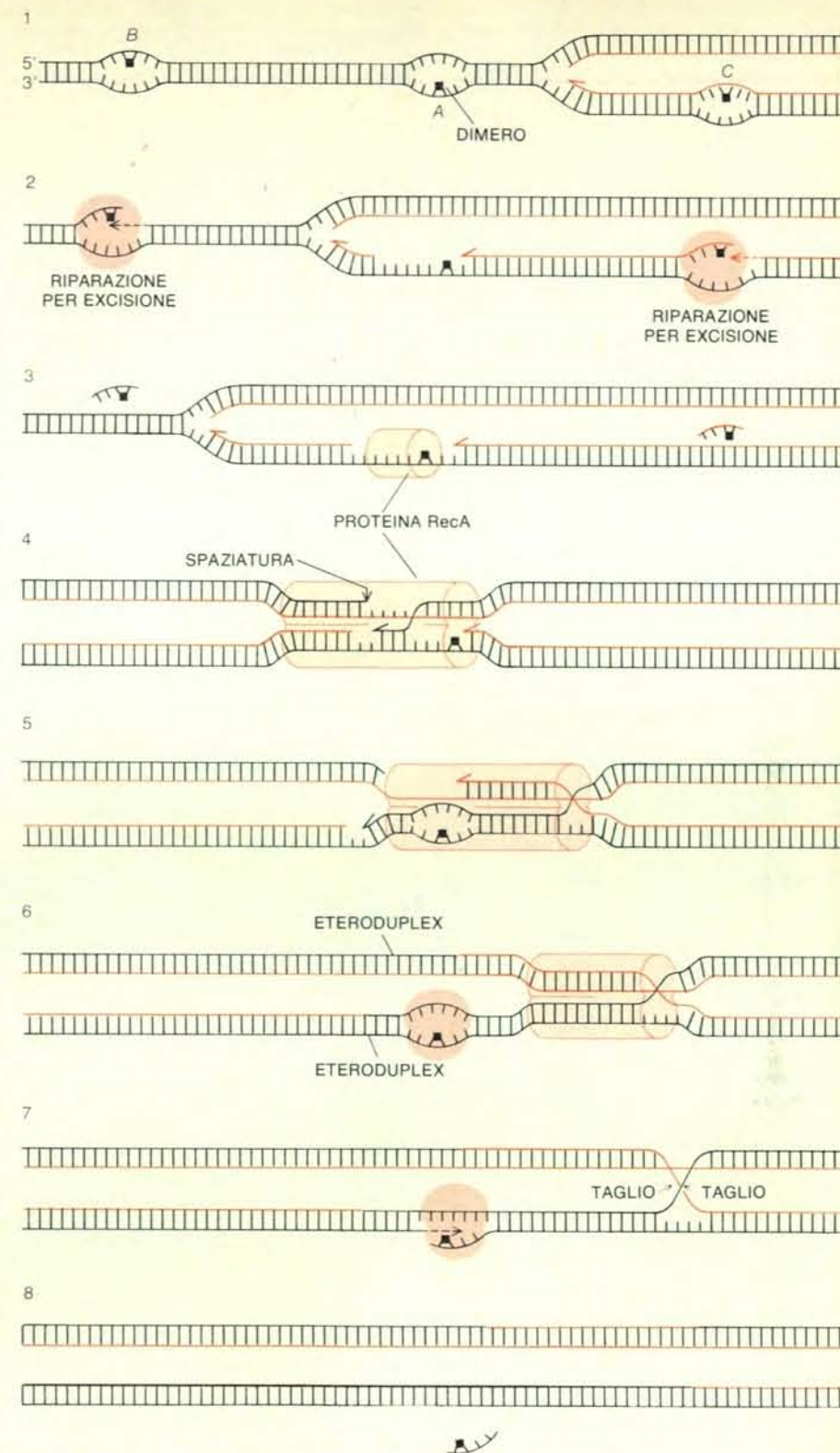
Il coinvolgimento di tale gene è ben superiore a questo. Alla metà degli anni settanta, fu dimostrato che gli stessi eventi che, in una cellula irradiata, fanno precipitare l'induzione lisogena, innescano la sintesi degli enzimi di riparazione del DNA; inoltre, interrompono la replicazione del DNA e la divisione cellulare, cosicché la cellula si allunga invece di dividersi in due cellule figlie, e fanno aumentare il tasso di mutagenesi, promuovendo, infine, la sintesi di notevoli quantitativi di una particolare proteina, la quale per un certo tempo fu conosciuta come proteina *X*. Queste risposte (e parecchie altre che non ho elencato qui) risultano essere delle reazioni a una situazione di emergenza; Miroslav Radman dell'Université Libre di Bruxelles diede a esse, collettivamente, il nome di «risposte di tipo *SOS*».

Nel 1975, Jacqueline George e Marc Castellazzi dell'Institut pour les Recherches de Biologie Moléculaire di Parigi trovarono che una particolare mutazione a carico del gene *recA* funge da interruttore che innesca la risposta di tipo *SOS* anche in assenza di un trattamento con raggi ultravioletti. George e Castellazzi lavoravano su un ceppo di *E.coli* lisogeno, caratterizzato da una mutazione termosensibile, la *recA441*, nota per essere letale per i batteri a una temperatura di 42 gradi centigradi. Notarono che la morte delle cellule batteriche poteva essere evitata se subentrava una seconda mutazione a carico dell'uno o l'altro di due altri geni, il *sulA* e il *sulB*. Ora che le cellule con la mutazione *recA441* potevano essere tenute in vita a 42 gradi, George e Castellazzi dimostrarono che, anche in assenza di radiazioni ultraviolette, tutte le funzioni di tipo *SOS* venivano innescate a quella temperatura. Se una mutazione nel gene *recA* poteva scatenare una risposta di tipo *SOS*, questa risposta doveva dipendere (almeno in parte) dal gene *recA*.

All'incirca nello stesso periodo, Jeffrey W. Roberts della Cornell University stava indagando sul meccanismo mediante il quale le radiazioni ultraviolette realizzano l'induzione del profago. Si sapeva che il profago lambda viene mantenuto nello stato quiescente da una proteina regolatrice, detta repressore, che mantiene inattivi tutti i geni del fago (tranne quello che specifica lo stesso repressore). Il repressore si lega alle regioni del DNA del fago note come «operatori» e impedisce il legame della RNA-polimerasi, che altrimenti trascriverebbe i geni del fago in RNA. Cercando il meccanismo mediante il quale le radiazioni ultraviolette liberano il repressore del fago lambda, Roberts e collaboratori trovarono che, nel corso dell'induzione, il repressore apparentemente scompariva. È poi risultato che esso viene distaccato dal DNA del profago perché è stato tagliato in due. I ricercatori della Cornell University hanno identificato e purificato l'enzima responsabile della scissione: una proteasi (enzima proteolitico) con un peso molecolare di 38 000. Essa è stata individuata in quantità crescenti in cellule di tipo selvatico dopo trattamento con radiazioni, ma non nei mutanti *recA* negativi. Potrebbe essere, pertanto, il prodotto del gene *recA*. (Maggiori quantità di proteasi non furono scoperte neppure nelle cellule *lexA* negative, per ragioni che verranno chiarite più avanti.)

Roberts e collaboratori trovarono che la proteasi, che avevano purificato dai mutanti *recA441* (quelli in cui le funzioni di tipo *SOS* sono sempre presenti e attive a una temperatura di 42 gradi centigradi), era più attiva di quella ottenuta da cellule di tipo selvatico; d'altra parte, un mutante in cui l'induzione del profago non ha luogo anche dopo irradiazione non ha mostrato alcuna attività di scissione sul repressore. Procedendo con l'analisi, si vide che la proteasi era di tipo insolito: aveva un'azione molto lenta, i suoi bersagli specifici erano dei repressori (ad esempio il repressore del fago lambda) e la sua attività dipendeva dalla presenza non solo dell'ATP (la fonte primaria dell'energia biochimica), ma anche del DNA a singolo filamento. Roberts sospettava che la sua proteasi fosse un prodotto del gene *recA*, ma era difficile dimostrare che non si trattava soltanto di un componente minore del preparato fino a quando non fu possibile raggiungere un più elevato grado di purificazione.

L'identificazione finale del prodotto del gene *recA* e la sua purificazione in forte quantità rappresentarono un importante progresso, che venne conseguito da Roberts e collaboratori e da un certo numero di altri ricercatori: Peter T. Emmerson e Stephen C. West dell'Università di Newcastle upon Tyne, Lorraine J. Gudas e David W. Mount dell'Università dell'Arizona, John W. Little e D. G. Kleid dello Stanford Research Institute, Kevin McEntee dell'Università di Chicago e Hideyuki Ogawa e Tomoko Ogawa dell'Università di Osaka. Nel 1977 venne stabilito che la proteina *X*, che nel 1971 Masayori Inouye e Arthur B. Pardee della Princeton University avevano osservato nelle cellule irradiate, e che Gudas aveva trovato mancante o alterata nei mutanti *lexA* negativi e *recA* negativi, era di fatto la proteina RecA. Era stata così dischiusa la via agli studi biochimici su questa proteina e alle ricerche sul meccanismo della riparazione per postreplicazione.

Nel frattempo l'altra metà di quello che cominciava a essere conosciuto come «fenomeno *rec-lex*» era sotto esame. Nel 1973, Mount all'Università dell'Arizona e K. Brooks Low alla Yale University analizzarono un ceppo diploide parziale di *E.coli*: cellule che contenevano un breve segmento supplementare del DNA cromosomico, cosicché su questo cromosoma erano presenti due versioni di alcuni geni. Essi osservarono che i mutanti *lexA* negativi sono dominanti da un punto di vista genetico, cioè una cellula, che porta sia il gene di tipo selvatico *lexA* sia il gene mutante *lexA* negativo, esibisce le caratteristiche della cellula mutante, inclusa la sensibilità ai danni da raggi X. Ora, la dominanza in un gene mutante è spesso un segno che quel gene ha a che fare con la regolazione genetica. Per esempio, certe mutazioni nel gene *lacI* di *E.coli*, che specifica il repressore che controlla l'e-

La riparazione per postreplicazione comporta la partecipazione di un dimero pirimidinico *(A)*, che interferisce con la replicazione, durante la quale i filamenti progenitori *(in nero)* si srotolano in corrispondenza della biforcazione per la replicazione. Vengono così sintetizzati *(1)* due filamenti figli *(in colore)*; altri dimeri *(B, C)* vengono invece coinvolti nella riparazione per excisione. Il dimero *A* impedisce l'appaiamento delle basi lungo un tratto di un filamento progenitore, producendo una interruzione da postreplicazione, di fronte a un segmento di DNA a filamento singolo *(2)*. La proteina RecA *(in colore chiaro)* si lega alla regione che presenta un filamento singolo *(3)* e la allinea con una omologa del duplex fratello; quando si verifica l'appaiamento omologo, un enzima spazia il duplex *(4)*. La proteina RecA, inoltre, sospinge l'estremità libera del filamento progenitore del duplex nella interruzione, producendo così uno scambio tra filamenti incrociati *(5)*. L'eteroduplex superiore può, a questo punto, essere riparato dalla DNA-polimerasi. Con la sequenza corretta a posto, di fronte al dimero *A*, e con la proteina RecA libera, *A* viene sottoposto all'azione degli enzimi di riparazione per excisione *(6)*. Infine, un enzima effettua due tagli in corrispondenza del sito dove avviene lo scambio tra filamenti incrociati *(7)*, risolvendo il processo di ricombinazione e producendo due molecole intatte di eteroduplex *(8)*.

spressione dei geni interessati nel metabolismo del lattosio, sono dominanti. Sembrò possibile, per analogia, che anche il *lexA* codificasse per una proteina repressore in grado di regolare certi enzimi incaricati del processo di riparazione del DNA.

Mount pensava che i mutanti *lexA* negativi potessero essere sensibili alle radiazioni in quanto producono un repressore difettoso, cioè un repressore che non si stacca in seguito a irradiazione. (Ciò spiegherebbe la dominanza dei *lexA* negativi. Anche se in una cellula fosse presente il repressore normale, che può essere messo in libertà, la proteina difettosa *lexA* negativa continuerebbe a legarsi e a reprimere gli enzimi di riparazione.) Lavorando con un ceppo che portava una particolare mutazione *lexA* negativa, chiamata *lexA3*, Mount cercò delle cellule che erano ridiventate resistenti alle radiazioni ultraviolette. Ne trovò una che portava ancora la mutazione originaria e, inoltre, una nuova mutazione che designò *lexA51* e che localizzò vicino al sito *lexA*, o addirittura al suo interno. Chiaramente una simile mutazione influenzava la funzione governata da *lexA*, in modo tale che i geni, in precedenza repressi dalla proteina *lexA*, venivano lasciati liberi di agire, mentre venivano sintetizzati gli enzimi di riparazione e veniva riparato il danno provocato dalle radiazioni ultraviolette. Risultò così che i mutanti *lexA51* sintetizzano il repressore che non riesce però a legarsi correttamente con il DNA. Di conseguenza, non solo la sintesi di un enzima di riparazione, ma anche tutte le altre funzioni di tipo *SOS*, compresa la sintesi maggiorata di RecA, sono attive tutto il tempo, anche in assenza di radiazioni ultraviolette o di altri agenti che danneggiano il DNA. Così, si è riusciti a dimostrare che la risposta *SOS* dipende non solo dal gene *recA*, ma anche dal gene *lexA* e sembra probabile che il prodotto di quest'ultimo sia una proteina repressore, che controlla tutti i geni per la risposta di tipo *SOS*.

La fase successiva consistette nell'identificare la proteina *lexA*. Little (che si trovava a quell'epoca all'Università dell'Arizona) e Mount vi riuscirono utilizzando le tecniche del DNA ricombinante. Isolarono il gene *lexA*, lo tagliarono e lo saldarono a un plasmide (un piccolo frammento circolare di DNA non cromosomico, che si replica autonomamente nella cellula batterica); inserirono il plasmide ricombinante in *E.coli* e furono in grado di identificare e purificare la LexA, una proteina con peso molecolare 24 000. In vista del possibile ruolo di repressore della LexA, essi la sottoposero a una prova di suscettibilità alla scissione da parte della proteasi *recA* e fecero l'importante scoperta che la proteina RecA scinde la LexA e procede più rapidamente di quando scinde il repressore del fago lambda, inducendo a credere che la LexA sia il substrato primario della proteasi. Mark Ptashne e Roger Brent della Harvard University dimostrarono, in seguito, che il gene *lexA* è a sua volta represso dalla proteina LexA: cioè è represso dal suo stesso prodotto, come nel caso del gene per il repressore di lambda e per alcuni altri geni.

Se la proteina LexA è un repressore, su quali geni esercita la propria azione, oltre al proprio? Cynthia J. Kenyon e Graham C. Walker del Massachusetts Institute of Technology idearono un ingegnoso esperimento per cercare di conoscere quali geni di *E.coli* vengono innescati come risposta al danno provocato al loro DNA. Lavorando con un ceppo da cui erano stati eliminati tutti i geni che governano il metabolismo del lattosio, utilizzarono un metodo messo a punto da Malcolm J. Casadaban e da Stanley N. Cohen della Stanford University School of Medicine per inserire il gene *lacZ* in posizioni casuali lungo il cromosoma. Quindi cercarono delle cellule in grado di metabolizzare il lattosio solo se era presente un agente che danneggiava il DNA (come la mitomicina *C*). Poterono così avanzare l'ipotesi che, in simili cellule, in risposta al danno al DNA, qualche gene, normalmente represso, era stato innescato. La funzione normale del gene era stata, tuttavia, disturbata dall'inserimento del gene *lacZ*, ed era quest'ultimo che veniva espresso. Quando i geni così innescati furono identificati (in base alla loro incapacità di svolgere la loro funzione normale), risultò che tra di essi erano presenti i geni *uvrA* e *uvrB* per gli enzimi di riparazione e, inoltre, il gene *umuC*, che è coinvolto nell'accresciuta mutagenesi conseguente a un trattamento con raggi ultravioletti.

L'esperimento eseguito al MIT aveva identificato certi geni che, chiaramente, erano repressi dalla proteina LexA. Barry M. Kacinsky, Aziz Sancar e Rupp cercarono una prova diretta di una simile regolazione e trovarono che il mutante *lexA51* di Mount, in cui sono sempre innescate le funzioni di tipo *SOS*, in quanto non vi è alcuna repressione da parte della proteina LexA, mostra un'attività *uvrA* e *uvrB* nell'incidere il DNA danneggiato di gran lunga superiore a quella delle cellule normali. Questa maggiore attività stimola anche una riparazione del DNA: Jeffrey Auerbach, nel mio laboratorio, ha individuato un maggior livello di sopravvivenza cellulare, che si potrebbe far risalire alla maggiore attività *uvrA* e *uvrB*, nei mutanti *lexA51* rispetto alle cellule di tipo selvatico.

Se la proteina LexA è realmente un repressore di un gruppo di geni interessati nella riparazione del DNA, deve legarsi agli operatori, vicino all'inizio di ciascuno di questi geni: ci si aspetterebbe di trovare, in tali regioni, sequenze nucleotidiche simili. Finora la sequenza del gene *recA* è stata determinata da Toshihiro Horii e dagli Ogawa a Osaka e, inoltre, da Sancar e Rupp; la sequenza dell'operatore del *lexA* dal gruppo di Osaka e da Little e Mount, e parti dei geni *uvrA* e *uvrB* da Sancar e Rupp. Una sequenza molto simile, lunga circa 20 nucleotidi, si trova, in effetti, nelle regioni degli operatori di ogni gene e questa «cassetta di pronto soccorso» ha (in varia misura in differenti geni) la forma ripetitiva invertita che è tipica dei siti di legame per il repressore. Nel caso dello stesso gene *lexA* vi sono, in realtà, due sequenze adiacenti, quasi identiche, che sono analoghe ai siti di legame multipli, presenti nell'operatore del fago lambda. Di recente, la proteina LexA è stata purificata da Little e Mount e da Brent e Ptashne e si è visto che si lega in maniera specifica alle «cassette di pronto soccorso» dei geni *recA* e *lexA*.

Gli eventi correlati della risposta *SOS* e la loro regolazione possono ora essere descritti ad ampie linee. In una cellula in fase di accrescimento e non danneggiata, i geni per la riparazione del DNA sono quasi completamente repressi dalla proteina LexA; tuttavia, viene sintetizzata una sufficiente quantità di repressore per occupare le regioni degli operatori, e una sufficiente quantità di proteine *uvr* per far fronte a sporadiche riparazioni per excisione. Quando le cellule sono esposte a una dose significativa di radiazioni ultraviolette, i dimeri pirimidinici, a cui non è possibile far fronte con le poche molecole di enzimi per l'excisione presenti nello stato non indotto, danno origine a interruzioni da postreplicazione. Qualunque proteina RecA già presente nella cellula si lega al filamento singolo di DNA che si trova di fronte a tali interruzioni; il DNA stimola l'attività proteolitica della RecA e viene di conseguenza indotta la risposta di tipo *SOS*. Tale proteina scinde i repressori LexA, liberando il gene *recA* e gli altri geni regolati dalla repressione; vengono sintetizzati forti quantitativi di proteina RecA e la quantità in eccesso si lega al DNA a filamento singolo ancora libero e quindi alle regioni adiacenti, a filamento doppio. In queste condizioni, la proteina LexA viene scissa non appena sintetizzata e i geni regolati rimangono innescati fino a quando il DNA a filamento singolo è presente nella cellula per stimolare l'attività proteasica della RecA. Dopo il completamento della riparazione per postreplicazione, con tutto il DNA ormai in

recA

5'···T A C T G T A T G A G C A T A C A G T A···
3'···A T G A C A T A C T C G T A T G T C A T···

lexA

5'···T G C T G T A T A T A C T C A C A G C A T A A C T G T A T A T A C A C C C A G G G···
3'···A C G A C A T A T A T G A G T G T C G T A T T G A C A T A T A T G T G G G T C C C···

uvrA

5'···T A C T G T A T A T T C A T T C A G G T···
3'···A T G A C A T A T A A G T A A G T C C A···

uvrB

5'···A A C T G T T T T T T T A T C C A G T A···
3'···T T G A C A A A A A A A T A G G T C A T···

Le regioni che fungono da operatore vicino all'inizio di ciascuno di quattro geni, controllati dalla proteina repressore LexA, hanno sequenze di basi simili, della lunghezza di circa 20 coppie: sono i siti di legame per il repressore. (Il legame è stato dimostrato finora solo per le sequenze dei geni *recA* e *lexA*.) In questo schema, le coppie di basi negli operatori *lexA*, *uvrA* e *uvrB*, che sono identiche a quelle nell'operatore *recA*, appaiono in colore, mentre le coppie di basi che si diversificano sono indicate in nero. Tutti i siti di legame, detti «cassette di pronto soccorso», comprendono sequenze ripetitive invertite (*frecce*), che leggono indifferentemente su ambedue i filamenti nella direzione 5'-3'. Includono le sequenze comuni *CTG* e *CAG* (*riquadri in colore*), separate da dieci basi. Il gene *lexA* ha non una, ma due cassette di pronto soccorso quasi uguali.

Il sistema di regolazione, basato sui geni *lexA* e *recA*, è quiescente durante l'accrescimento normale in assenza di danni al DNA (*in alto*). Il repressore LexA si lega agli operatori dei geni *lexA*, *recA*, *uvrA*, *uvrB* e di alcuni altri geni, mantenendo la sintesi dell'RNA-messaggero e della proteina al basso livello caratteristico delle cellule non indotte. Un danno al DNA, sufficiente a produrre una interruzione da postreplicazione, attiva la risposta di tipo *SOS* (*in basso*). La proteina RecA si lega al DNA a filamento singolo, che si trova di fronte alla interruzione; la sua attività proteolitica viene pertanto stimolata e il repressore LexA viene scisso. In mancanza di repressore funzionale, i geni controllati dal LexA si attivano e la proteina viene sintetizzata con una maggiore velocità. (Anche se viene sintetizzata più LexA, l'eccesso è scisso quando la proteina RecA è attivata dal DNA a filamento singolo.) Quando la riparazione del DNA è giunta a conclusione, la RecA non scinde più il repressore; il repressore LexA appena sintetizzato si lega di nuovo con gli operatori e la cellula ritorna allo stato non indotto.

La metilazione di un sito di una molecola di guanina da parte dell'agente chimico *MNNG* viene annullata da un successivo processo che sembra costituire un tipico esempio di risposta adattativa al danno provocato al DNA. Il *MNNG* aggiunge un gruppo metilico (CH_3) a vari siti del DNA, tra cui un atomo di ossigeno nella posizione 6 della guanina *(disegno a sinistra)*, rompendo così il legame che questa ha con la citosina. La riparazione viene qui realizzata da una proteina che è sintetizzata nella cellula durante un periodo di adattamento, cioè quando la cellula stessa si trova esposta a una bassa concentrazione di *MNNG*. Un residuo di cisteina (uno dei venti amminoacidi) nella proteina agisce da accettore di metile: si lega infatti al gruppo CH_3, riportando in questo modo la guanina alla sua condizione originaria *(disegno a destra)*.

forma di doppio filamento, la RecA non viene più attivata come proteasi. La proteina LexA sintetizzata di recente non viene più scissa e può funzionare da repressore; la cellula ritorna allo stato non indotto.

Chiaramente la risposta di tipo *SOS* promuove una efficiente riparazione del DNA, facendo aumentare la disponibilità degli enzimi responsabili della riparazione per excisione e aprendo così la via alla riparazione per postreplicazione. Nella riparazione per excisione, gli enzimi specificati dai geni *uvrA*, *uvrB* e *uvrC* agiscono di concerto per riconoscere le basi danneggiate e per avviare il processo di excisione. Dato che le cellule riescono a sopravvivere a molte migliaia di queste lesioni anche se il processo è lento, sono necessari forti quantitativi dei suddetti enzimi. Rupp e collaboratori hanno dimostrato che una cellula non danneggiata possiede normalmente solo 10 o 20 molecole dell'enzima specificato dal gene *uvrA*: l'incremento dell'attività di sintesi a seguito dell'induzione *SOS* rende disponibile una sufficiente quantità di enzima per una excisione efficiente dei dimeri in grande numero. Il preciso meccanismo d'azione dei tre enzimi è oggi sotto esame in parecchi laboratori. I geni per gli enzimi sono stati clonati su plasmidi a copie multiple per aumentarne i livelli di sintesi. Inoltre, l'efficace metodo della «maxicellula», messo a punto da Sancar e Rupp, permette di identificare e di purificare gli enzimi anche se non sono ancora disponibili pratici saggi funzionali per le proteine.

Come il lettore avrà notato, la RecA ha due funzioni molto diverse: agisce come proteasi nel suo ruolo di regolatore così come enzima della ricombinazione, che manipola filamenti di DNA. Nella ricombinazione, quale si verifica nella riparazione per postreplicazione, la sua prima funzione è di promuovere un appaiamento omologo tra due molecole di DNA, cioè nel portare l'una accanto all'altra, in un allineamento preciso, regioni aventi una sequenza di basi complementari, in vista di uno scambio di filamenti. Nel 1979, McEntee di Chicago, George M. Weinstock e I. Robert Lehman della Stanford School of Medicine e Takahiko Shibata, Chanchal DasGupta, Richard P. Cunningham e Charles M. Radding della Yale University hanno dimostrato che la proteina RecA può promuovere un analogo appaiamento di basi tra un singolo filamento di DNA e un duplex. Il gruppo di Radding ed Era Cassuto e West nel mio laboratorio hanno trovato che essa può promuovere un appaiamento omologo anche tra due molecole circolari duplex, se una delle due ha una regione a filamento singolo equivalente a una interruzione postreplicativa, ma hanno trovato anche che un simile appaiamento non viene promosso tra due duplex integri.

L'esatto meccanismo dell'appaiamento deve ancora essere determinato, ma sembra probabile che il duplex con una interruzione a filamento singolo e il duplex integro siano portati, in primo luogo, dalla RecA a un contatto non specifico e quindi spostati l'uno rispetto all'altro fino a quando le loro basi complementari non sono allineate. Una certa comprensione del processo si può acquisire con lo studio dei suoi fabbisogni energetici. L'energia viene chiaramente fornita dall'ATP e troviamo che la demolizione dell'ATP (per estrarre la sua energia) da parte della RecA viene stimolata sia dalle regioni a filamento singolo sia dai segmenti duplex confinanti, che presentano interruzioni a filamento singolo; ma non dai duplex integri. Chiaramente la proteina adotta una configurazione attiva nei duplex che presentano interruzioni a filamento singolo. Forti quantitativi di proteina RecA (una molecola ogni cinque coppie di basi) sono indispensabili per un appaiamento omologo, sufficiente a dare al filamento, che presenta una interruzione, una struttura filamentosa di sostegno che porta le molecole di DNA a contatto e le fa spostare l'una di seguito all'altra fino a quando non si stabiliscono dei contatti complementari. Il movimento relativo ricorda quello di un bruco, con il DNA che presenta interruzioni che si allunga e poi si contrae, allorché la proteina RecA si lega con esso e quindi si stacca nel momento in cui l'ATP viene demolito.

La RecA, avendo fatto allineare le molecole di DNA in una particolare regione, promuove l'effettivo scambio di filamenti. West, lavorando con molecole duplex dotate di adatte regioni a filamento singolo, ha dimostrato che la RecA effettua uno scambio reciproco tra due regioni a duplice filamento; ne risultano due eteroduplex in cui i quattro filamenti si sono scambiati il partner. Ciò può avvenire grazie a uno scambio tra filamenti singoli, proposto come struttura intermedia nella ricombinazione da Robin Holliday del National Institute for Medical Research di Londra nell'anno 1965. In presenza di ATP, la proteina RecA dirige l'effettivo scambio dei filamenti, a mano a mano che il sito di scambio procede sulle molecole. Questo processo prende il nome di migrazione ramificata. Sono quindi necessari, per completare il processo di riparazione, enzimi diversi dalla RecA: una nucleasi per tagliare i filamenti in vari stadi dello scambio, la DNA-polimerasi per inserire le basi mancanti per colmare le interruzioni, una ligasi per saldare l'impalcatura zucchero-fosfato e forse altri enzimi ancora.

Ho già ricordato che una più intensa mutagenesi sembra essere un aspetto della risposta di tipo *SOS*. Alcuni anni fa, Jean J. Weigle del California Institute of Technology, studiando nel fago lambda la mutagenesi indotta da radiazioni ultraviolette, scoprì che sia la frequenza di mutazione, sia la sopravvivenza dei fagi irradiati aumentavano quando questi fagi infettavano cellule ospiti che erano state irradiate circa 30 minuti prima. Evelyn M. Witkin della Rutgers University e Junichi Tomizawa, che è oggi al National Institute of Arthritis, Metabolism and Digestive Diseases, hanno trovato che la mutazione non viene indotta dalle radiazioni nei batteri *recA* o *lexA* negativi.

La Witkin ha suggerito che una speciale «riparazione soggetta a errori» è responsabile della maggior sopravvivenza, associata alla mutagenesi, nei batteri irradiati di tipo selvatico. Con Radman essa ha ipotizzato che la risposta *SOS* promuova la comparsa di una speciale DNA-polimerasi, che tollera più dell'ordinario mutazioni nel DNA che funge da stampo e può replicare tale DNA, osservando meno rigorosamente della comune polimerasi le regole dell'appaiamento delle basi. Una polimerasi soggetta a errori potrebbe far aumentare il livello di sopravvivenza a spese di una maggiore frequenza di mutazione. Proprio di recente, Stuart M. Linn, David Lackey e Sharon W. Krauss di Berkeley hanno rilevato piccolissime quantità della DNA-polimerasi soggetta a errori in cellule in cui era stata indotta una maggior sintesi di proteina RecA. La mutagenesi in generale e la mutagenesi da radiazioni ultraviolette, in particolare, comportano meccanismi di elevata complessità, che solo ora cominciamo a poter spiegare.

La riparazione per excisione e quella per postreplicazione sembrano essere meccanismi fondamentali per fronteggiare il danno provocato al DNA, ma per particolari lesioni vi sono rimedi più specifici. Per esempio, singole basi non giuste o alterate vengono rimosse da un certo numero di enzimi, chiamati glicosilasi, estremamente specifici per particolari tipi di basi lese.

Un sistema di riparazione scoperto di recente è particolarmente interessante sia a causa del suo insolito meccanismo, sia perché sembra rappresenti un secondo tipo di risposta indotta al danno genetico, ben distinto dalla risposta di tipo *SOS*. Si chiama, infatti, «risposta adattativa» e mitiga il danno al DNA che deriva dalla metilazione (aggiunta di un gruppo CH_3) di una base guanina.

La risposta adattativa venne scoperta nel 1977 da Leona Samson e da John Cairns, che si trovavano allora agli Imperial Cancer Research Fund Laboratories di Londra ed erano interessati ai possibili effetti di una lunga esposizione a livelli molto bassi di sostanze chimiche, potenzialmente cancerogene. Essi studiavano su *E. coli* l'effetto di esposizioni prolungate a basse concentrazioni di un potente mutageno usato in laboratorio, la *N*-metil-*N'*-nitronitrosoguanidina *(MNNG)*. Con loro sorpresa, due ore di incubazione in un solo microgrammo di questa sostanza per millilitro di mezzo culturale rendevano i batteri estremamente resistenti a successive esposizioni al farmaco in concentrazioni centinaia di volte superiori. Questa risposta adattativa, che venne fatta risalire a un gene designato *ada*, durava un'ora o due, periodo in cui il livello sia della mutagenesi sia della morte cellulare derivante da metilazione è di gran lunga inferiore a quello che si riscontra in cellule che non hanno subito adattamenti. La suscettibilità al danno da radiazioni ultraviolette delle cellule adattate non viene influenzata in modo analogo, e per questo motivo la risposta adattativa non può essere semplicemente un aspetto della risposta di tipo *SOS*.

Thomas Lindahl e collaboratori, della Università di Göteborg in Svezia, hanno analizzato il meccanismo biochimico che è alla base della risposta adattativa. La *MNNG* metila le basi in parecchi punti. La risposta adattativa sembra essere indotta quando viene coinvolto un atomo di ossigeno (designato come O^6) della guanina. Il gruppo metilico inserito sporge su quello che viene considerato il principale solco della doppia elica del DNA. Mentre la maggior parte delle basi alterate si stacca o viene excisa dal DNA per opera di opportuni enzimi di riparazione e quindi sostituita, Lindahl ha trovato che, in questo caso, la guanina rimane al suo posto e viene rimosso il gruppo metilico lesivo. Il compito viene assolto da una speciale proteina, sintetizzata in abbondanza solo durante la risposta adattativa e che funge da accettore di gruppi metilici. Un carattere distintivo del processo è la sua capacità finita. Quando una unità amminoacidica (per esempio, una molecola di cisteina) della proteina di riparazione ha catturato un gruppo metilico (formando una *S*-metilcisteina), la capacità di riparazione della molecola si esaurisce.

Devono ancora essere determinate le ramificazioni complete della risposta adattativa. Molti aspetti della enzimologia della risposta di tipo *SOS* e l'intera storia del suo controllo da parte della LexA devono ancora essere chiariti nei particolari. Sicuramente verranno scoperti altri sistemi indotti di riparazione. Il loro studio nei batteri è ancora ben lungi dall'essere completo e viene realizzato oggi in molti laboratori. Nel contempo, molte delle attuali ricerche si incentrano sugli enzimi di riparazione del DNA e sulla risposta al danno genetico in organismi superiori come i mammiferi, motivate in parte dalla convinzione sempre più diffusa che il danno al DNA e la sua riparazione siano fattori significativi nella proliferazione cancerosa, in altre malattie e nel processo della senescenza.

SCIENZA IN CASA

di Jearl Walker

L'aerodinamica delle samare

Le samare sono frutti secchi e alati di frassino, olmo, acero e altri alberi, che cadono a terra volando in larghe spirali e trasportando un unico seme. Vedendoli volare viene spontaneo chiedersi come si comportino questi frutti (che la gente abitualmente chiama semi) dal punto di vista aerodinamico, ma la risposta non è facile.

Il primo studio dettagliato dell'aerodinamica di una samara con una sola ala, come quella dell'acero, è stato condotto da R. Åke Norberg, dell'Università di Göteborg. Altre ricerche sono state fatte da Charles W. McCutchen del National Institute of Arthritis, Metabolism and Digestive Diseases e da F. M. Burrows dell'University of North Wales. Darò qui un resoconto del lavoro di Norberg, che ha studiato la samara dal punto di vista dell'aerodinamica degli elicotteri, seguito da un brevissimo accenno alle ricerche di McCutchen su samare diverse da quelle dell'acero.

Norberg è partito da due assunti fondamentali. Il primo è che la samara sia assimilabile a un'ala piana. Il secondo è che la massa dell'ala sia concentrata lungo l'asse longitudinale che va dal seme all'estremità alare.

Una samara d'acero che cade a terra ruotando non segue un moto accelerato, sotto l'azione della gravità, a causa della forza aerodinamica esercitata sull'ala. Per spiegare con semplicità l'origine e l'azione di tale forza si può considerare che la rotazione della samara avvenga intorno al suo centro di massa, che cade lungo la verticale (in realtà il centro di rotazione non coincide esattamente con il centro di massa). L'ala giace praticamente su un piano orizzontale e, ruotando, copre un disco entro il quale la velocità dell'aria è uniforme. È possibile analizzare il volo da due punti di vista: quello usuale di un osservatore fermo che guarda la samara che cade e quello di un osservatore che immagina di cadere stando a cavallo della samara.

Secondo il primo punto di vista la samara cade entro una colonna d'aria ferma. Ogni strato orizzontale d'aria all'interno di tale colonna viene spinto verso il basso dalla samara che lo attraversa. La colonna d'aria sopra il frutto si muove quindi con una certa velocità verso il basso, minore di quella di caduta della samara stessa. L'aria che costituisce questa parte della colonna mette in evidenza gli effetti di una forza diretta verso il basso perché ha subito un'accelerazione dallo stato di riposo a uno stato di movimento diretto verso il terreno. La corrispondente reazione sulla samara è quindi rivolta verso l'alto.

Dal punto di vista dell'osservatore che cade insieme alla samara, la colonna d'aria sottostante si muove verso l'alto con una velocità uguale e contraria a quella di caduta vista dal precedente osservatore stazionario. Ogni strato d'aria che attraversa il disco generato dalla rotazione della samara subisce un'accelerazione verso il basso e quindi si muove verso l'alto con una velocità minore di quella alla quale viene incontrato. Ciò riflette l'effetto di una forza rivolta verso il basso che ha diminuito la velocità di ascesa della colonna d'aria. La corrispondente reazione sulla samara è diretta verso l'alto. In entrambi i casi si osserva quindi che esiste una forza diretta verso l'alto che agisce sulla samara.

Questa forza aerodinamica è uguale e contraria al peso del frutto, che quindi scende verso il terreno con velocità costante. Per l'osservatore che scende insieme alla samara la colonna d'aria sottostante si avvicina a velocità costante. La velocità di caduta è proporzionale al quadrato di un rapporto detto carico del disco. Esso è dato dal peso della samara diviso per l'area del disco spazzato dall'ala durante un giro completo.

Diverse samare di acero cadono con velocità differenti perché non possiedono tutte il medesimo carico del disco. Una ragione ovvia di ciò va ricercata nel diverso peso e nelle diverse dimensioni delle varie samare. Una ragione meno ovvia risiede nel fatto che l'ala non sta proprio su un piano orizzontale. L'angolo fra il piano dell'ala e l'orizzontale varia da samara a samara. Alcuni frutti scendono tenendo l'ala praticamente orizzontale, altri, invece, la tengono inclinata anche di 45 gradi e più rispetto al piano dell'orizzonte. Il carico del disco è evidentemente maggiore per le ali inclinate di quanto non sia per quelle orizzontali, con la conseguenza di una maggior velocità di discesa delle prime.

L'analisi della forza aerodinamica su un'ala di samara inclinata è più difficile per un'osservatore esterno di quanto non sia per l'osservatore che scende insieme al frutto. Considereremo quindi solo quest'ultimo punto di vista, descritto nella figura a pagina 53. Il movimento relativo dell'aria che incontra l'ala è diretto verso l'alto e verso il bordo d'uscita dell'ala stessa. È questo movimento che produce la portanza necessaria al bilanciamento del peso della samara.

Per studiare la forza aerodinamica sull'ala, somma della portanza, diretta verso l'alto, e della resistenza all'avanzamento, occorre stabilire innanzitutto alcune definizioni. L'asse verticale intorno al quale ruota la samara è detto asse di rotazione. Poniamo che esso passi per il centro di massa della samara, che è vicino all'estremità che contiene il seme. L'apertura alare è la lunghezza dell'asse lungo dell'ala, che va dal seme all'estremità dell'ala. La corda, più corta dell'apertura, è perpendicolare a questa e va dal bordo d'entrata al bordo d'uscita dell'ala. L'inclinazione dell'ala rispetto al piano orizzontale nella direzione dell'apertura è detto diedro.

La forza agente sull'ala può venire calcolata dividendo l'ala stessa in fettine parallele alla corda. Si calcola approssimativamente la forza agente su ogni fettina moltiplicando la sua superficie per il quadrato della velocità dell'aria che si muove dietro di essa. Sommando le forze calcolate per tutte le singole sezioni si ottiene una stima della forza complessiva agente sull'ala sufficiente per i nostri fini.

La velocità dell'aria che si muove verso l'alto dietro l'ala può venire divisa in due componenti, una perpendicolare all'apertura e una parallela a questa. Solo la componente perpendicolare contribuisce alla portanza e alla resistenza. Il movimento dell'aria parallelo alle corde dell'ala è più difficile da analizzare, poiché la velocità dell'aria dipende dalla distanza della corda considerata dall'asse di rotazione. L'aria che si muove lungo una corda vicina all'asse di rotazione ha una velocità relativamente bassa, dato che tale corda si muove lentamente per la rotazione dell'ala. L'aria che si muove lungo una corda più lontana dall'asse di rotazione ha una velocità più elevata, a parità di velocità di rotazione. Per analizzare gli effetti del movimento dell'aria lungo una corda prenderò in considerazione una porzione intermedia dell'ala.

L'aria investe questa sezione d'ala muovendosi sia in direzione verticale sia all'indietro, lungo una corda. Del movimento in direzione verticale ci interessa solo la componente perpendicolare al piano dell'ala. Il vettore velocità dell'aria sopra la sezione d'ala considerata giace, infatti, su una direzione intermedia fra quella della perpendicolare all'ala e quella della corda.

Se il contributo del movimento in direzione verticale è relativamente elevato,

Riprese stroboscopiche di samare in volo

cioè se la samara scende velocemente, il vettore velocità dell'aria è praticamente verticale. Se invece tale contributo è piccolo, la velocità ha direzione compresa fra la perpendicolare all'ala e la corda di questa. La direzione assunta dal vettore velocità è importante per la stabilità del movimento di una samara.

Il valore della velocità dell'aria determina il valore della forza sulla sezione d'ala considerata. Il maggior contributo alla forza aerodinamica totale sull'ala è dovuto alle sezioni d'ala più lontane dal centro di rotazione. Ciò è dovuto, in parte, al fatto che ivi la corda è più lunga e quindi la superficie delle sezioni è maggiore. Inoltre la velocità dell'aria è più grande, a causa del movimento rotatorio dell'ala. Quindi, se una samara deve avere una grande portanza, è giusto che la corda dell'ala aumenti con la distanza dal centro di rotazione.

La forza aerodinamica totale è perpendicolare all'ala e giace su un piano verticale che attraversa l'apertura. La componente verticale di tale forza bilancia il peso del seme. Se si lascia cadere una samara, questa deve in qualche modo orientarsi per bilanciare il suo peso con la componente verticale della forza aerodinamica. È interessante osservare che tale orientamento viene raggiunto automaticamente.

Ottenuto il giusto orientamento, questo deve venire mantenuto anche in presenza di perturbazioni prodotte dal vento. La maggior parte delle samare sono intrinsecamente stabili per almeno quattro diversi fattori concorrenti: l'angolo di incidenza dell'ala, il diedro, l'inclinazione sull'orizzontale del cerchio tracciato dall'estremità alare e la possibilità di movimenti laterali.

L'angolo di incidenza è quello fra la corda dell'ala e la direzione del vettore velocità dell'aria che la investe. La figura in basso a pagina 54 mostra una sezione alare di samara circa a metà dell'apertura. La forza aerodinamica che agisce su tale sezione è indicata da un vettore applicato in un punto che viene detto centro di spinta. La posizione di questo dipende dal profilo della sezione alare e dalla direzione dell'aria che la investe, cioè, appunto, dall'angolo di incidenza.

La sezione d'ala possiede anche un centro di massa, sul quale possiamo immaginare applicato il peso della sezione stessa. Il peso è però trascurabile in confronto alla forza centrifuga cui è soggetta la sezione. Questa forza, dovuta alla rotazione della samara, agisce anch'essa sul centro di massa.

L'orientamento della sezione rispetto alla velocità dell'aria che la investe è stabile se il centro di spinta coincide con il centro di massa. Quindi la stabilità richiede un ben determinato angolo d'incidenza. Poniamo che l'ala venga perturbata in modo che il centro di pressione si sposti verso il bordo d'uscita. Tale orientamento cabrato produce un aumento dell'angolo di incidenza e quindi modifica la forza aerodinamica agente sulla sezione. La sua componente verticale non riesce più a bilanciare il peso della samara, come è invece necessario per avere un volo stabile e prolungato.

La samara possiede però una forma che le consente di correggere la situazione riprendendo il giusto angolo d'incidenza. La forza che agisce sul centro di spinta produce un momento che fa ruotare la sezione intorno al centro di massa, ripristinandone l'orientamento iniziale. Il centro di spinta torna a coincidere con il centro di massa e la samara ridiventa stabile. Qualcosa di simile capita anche se la perturbazione fa diminuire l'angolo d'incidenza rispetto al suo valore corretto, spostando in avanti il centro di spinta e facendo picchiare la samara. Ancora una volta la forza gente sul centro di spinta fa ruotare la sezione intorno al suo centro di massa, facendole riprendere il giusto orientamento.

l'ala ruota intorno
al baricentro della samara,
che scende verticalmente

l'ala ruota intorno
al baricentro della
samara

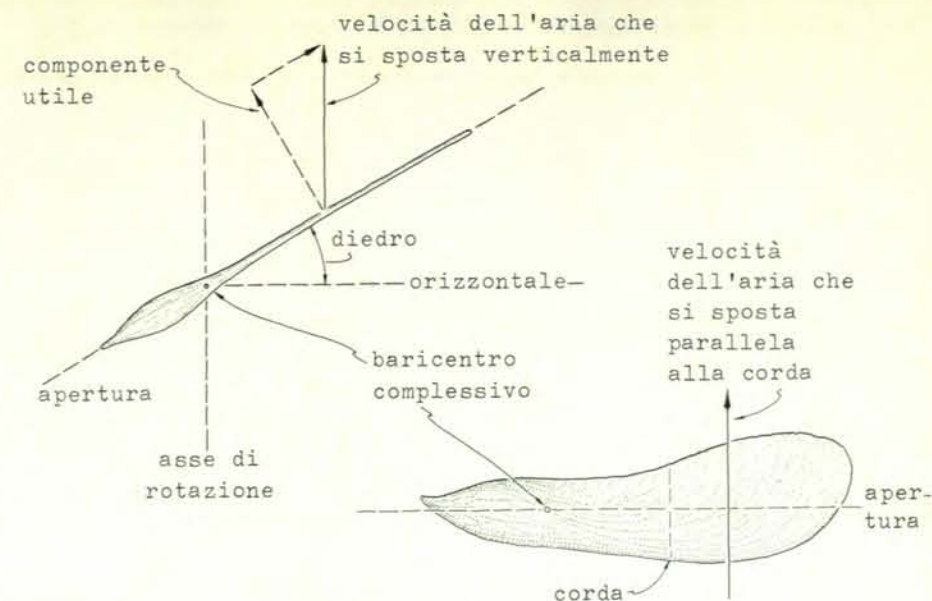
il baricentro si sposta
lungo una spirale con
verso di rotazione opposto

I possibili percorsi di volo di una samara d'acero

In base agli studi sull'aerodinamica dei profili alari piani, tale ricupero automatico del giusto angolo di incidenza è possibile solo per una ben determinata distribuzione dei pesi lungo la corda della sezione. Torniamo a considerare una sezione trasversale dell'ala. Il centro di massa di tale sezione deve trovarsi a una distanza dal bordo d'attacco compresa fra il 27 e il 35 per cento della lunghezza della corda di tale sezione. Il centro di massa deve anche trovarsi dietro alla posizione più avanzata raggiungibile dal centro di spinta, altrimenti l'ala non potrebbe recuperare da una picchiata. L'evoluzione ha fatto sì che la samara abbia il centro di massa in tale posizione. L'ala è infatti costituita da una nervatura anteriore che si appiattisce andando verso il bordo d'uscita. La massa della nervatura, concentrata in corrispondenza del bordo d'attacco, sposta in avanti il centro di massa.

È anche necessario che rimanga stabile l'angolo di planata, fra la direzione di volo e il piano orizzontale. La samara percorre una traiettoria a spirale, ma possiamo ancora definire un angolo di planata considerando la direzione della velocità istantanea di discesa. L'angolo di planata è importante perché da esso dipende l'orientamento della forza aerodinamica sull'ala. Quando esso coincide con il cosiddetto angolo naturale di planata, la forza aerodinamica si trova su un piano verticale che attraversa l'apertura alare. Se la planata è troppo picchiata, il vettore che indica la direzione della forza si inclina verso il bordo d'attacco dell'ala. Se la planata è invece eccessivamente tesa, la forza aerodinamica si inclina verso il bordo d'uscita. Nessuno di questi due orientamenti può consentire un volo stabile.

Non appena la samara inizia la sua caduta verso terra deve raggiungere rapidamente il giusto angolo di planata, in modo che la forza aerodinamica venga a trovarsi su un piano verticale. Solo allora la componente verticale della forza può controbilanciare il peso della samara, consentendo il raggiungimento di una velocità di discesa costante. Nelle prime fasi della discesa la componente verticale della forza è inferiore al peso, perché la velocità dell'aria è minore di quanto dovrebbe. La samara deve regolare in conseguenza il proprio angolo di planata, altrimenti cadrebbe con moto accelerato fino a raggiungere il terreno.

Le fasi della correzione sono le seguenti: dato che la samara accelera il movimento verso il basso, il vettore velocità

L'andamento della velocità dell'aria

dell'aria è più prossimo alla verticale di quanto dovrebbe; l'angolo di incidenza è quindi troppo grande e la samara picchia per diminuirlo. La forza aerodinamica si inclina verso il bordo d'entrata dell'ala e ciò fa aumentare la velocità di rotazione della samara.

Dal punto di vista dell'osservatore immaginario che cavalca la samara ciò fa aumentare la velocità dell'aria che si muove verso il bordo d'uscita parallelamente alla corda. Il vettore velocità risultante (somma della velocità dell'aria che si muove verticalmente in seguito alla caduta e di quella che si muove lungo la corda a causa della rotazione) era inizialmente troppo verticale. In seguito all'aumento della velocità dell'aria lungo la corda, il vettore si inclina verso il basso e all'indietro; in queste condizioni l'angolo d'incidenza è troppo piccolo e l'ala cabra. Ciò riporta la forza aerodinamica sul piano verticale e la velocità di rotazione cessa di aumentare.

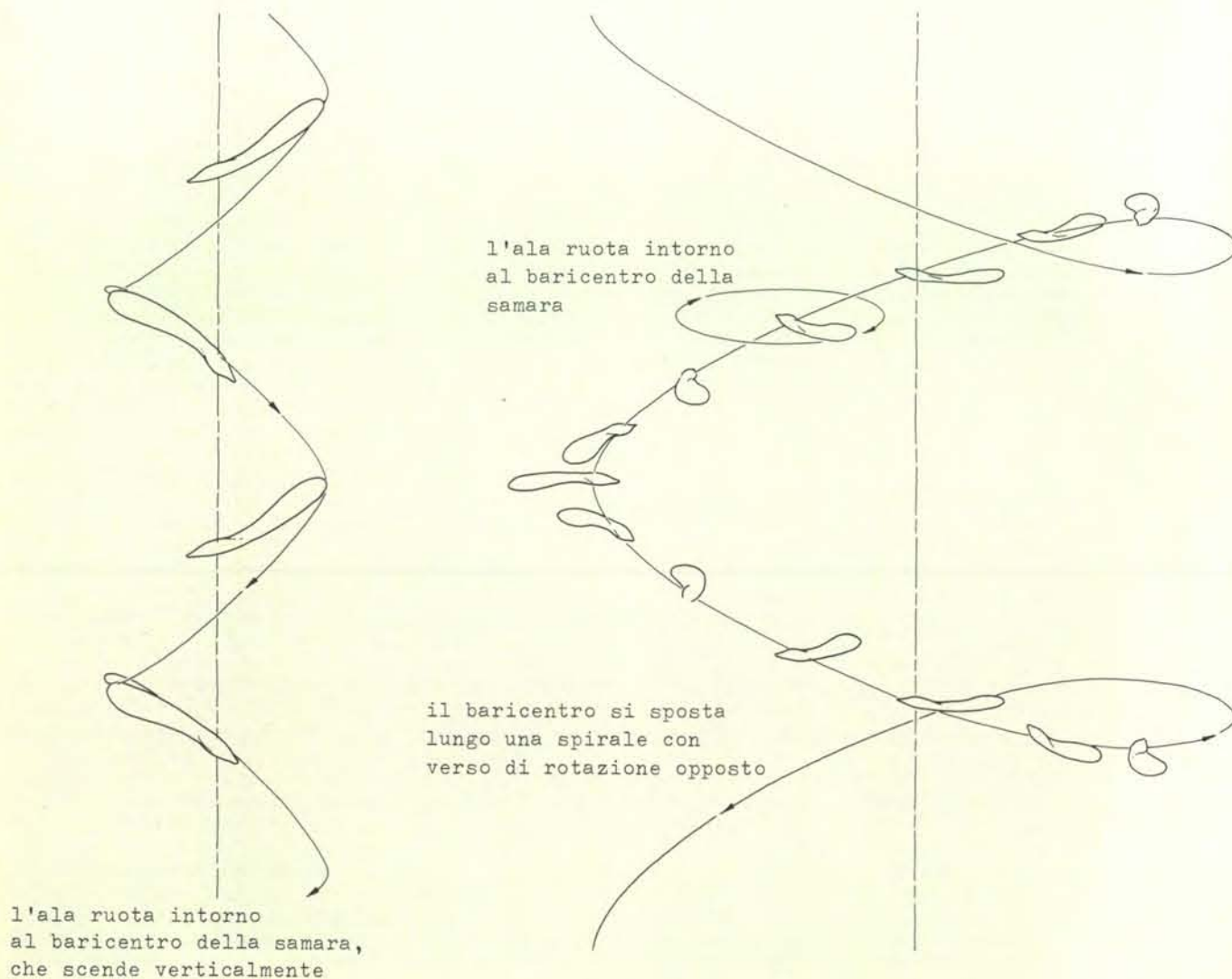
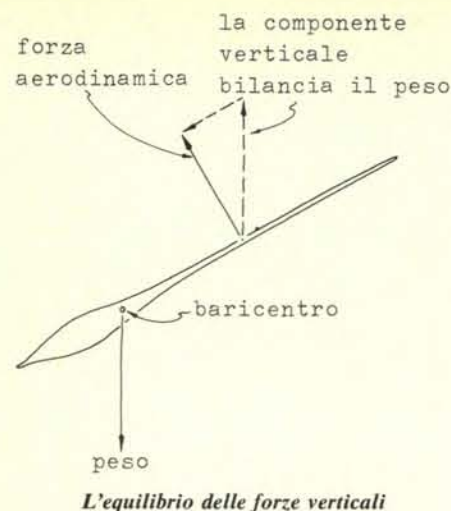
Nel corso della manovra di correzione la samara ha aumentato la propria velocità di discesa (perché inizialmente subiva un'accelerazione verso il basso) e la propria velocità di rotazione (perché l'inclinazione della forza aerodinamica durante le prime fasi della correzione spingeva l'ala in avanti intorno all'asse di rotazione). Ne segue che al termine della manovra la velocità dell'aria è aumentata e così pure la forza aerodinamica, che riesce ora a bilanciare il peso della samara. Da questo momento in poi il frutto discende con velocità di caduta e di rotazione costanti e con i giusti angoli di incidenza e di planata.

Se la traiettoria iniziale di discesa fa sì che la velocità dell'aria sia troppo elevata, la componente verticale della forza aerodinamica supera il peso. La samara deve eseguire una manovra di correzione per ridurre l'intensità della forza aerodinamica. Mentre continua la caduta della samara l'eccessiva forza aerodinamica ne riduce la velocità. Dal punto di vista dell'osservatore che sta sul frutto, tale diminuzione della velocità di discesa non risulta solo in una diminuzione della velocità dell'aria incidente, ma anche in una variazione dell'inclinazione del vettore velocità, che tende a spostarsi all'indietro, verso la direzione della corda. Ciò rende troppo piccolo l'angolo di incidenza e l'ala deve cabrare per riguadagnare l'angolo corretto.

Questo movimento sposta all'indietro la forza aerodinamica e tale inclinazione si oppone alla rotazione della samara, che rallenta. Per l'osservatore solidale con la samara, tale riduzione della velocità di rotazione corrisponde a una diminuzione della velocità dell'aria che si muove parallelamente alla corda. Ciò fa ancora diminuire la velocità risultante dell'aria incidente, mentre cambia anche la sua direzione, ora più vicina alla verticale. L'angolo di incidenza risulta ora eccessivo e la samara deve picchiare per ridurlo. Questa manovra riporta la forza su un piano verticale che attraversa l'apertura e quindi fa cessare la riduzione della velocità di rotazione.

Ne consegue una riduzione della velocità di discesa (a causa dell'eccessiva forza aerodinamica presente all'inizio) e della velocità di rotazione (in conseguenza dell'inclinazione all'indietro della forza durante le prime fasi dell'aggiustamento). Il vettore velocità dell'aria è ora più piccolo, e così pure la forza aerodinamica. La componente verticale di questa è ora pari al peso della samara, con la conseguenza che il seme volante scende a terra con velocità di avanzamento e di rotazione costanti e con i corretti angoli di incidenza e di planata.

Una perturbazione casuale dell'angolo di planata viene corretta con meccanismi simili. Per esempio, se la samara improvvisamente si mette a picchiare, aumen-

L'equilibrio delle forze verticali

tando così l'angolo di planata, il vettore velocità dell'aria si sposta troppo vicino alla verticale. Dato che l'angolo di incidenza cessa allora di essere corretto, la samara picchia ancora di più per correggerlo. Il vettore forza aerodinamica si inclina in avanti e la velocità di rotazione aumenta riportando verso l'orizzontale il vettore velocità dell'aria. Deve ora venire una seconda correzione dell'angolo di incidenza. Questa volta l'ala cabra riportando verso la verticale il vettore forza. A questo punto la samara dovrebbe avere riguadagnati i giusti angoli di incidenza e di planata. È però facile che la correzione sia eccessiva e che le occorra tutta una serie di correzioni di ampiezza decrescente per raggiungere una condizione di planata stabile.

Anche il diedro dell'ala viene riportato al giusto valore dopo piccole perturbazioni. Questo angolo è complementare all'apertura del cono che verrebbe generato dall'ala durante la rotazione se la samara non si spostasse in basso durante la rotazione (la superficie generata nel moto rototraslatorio dell'ala assomiglia a certi trucioli tagliati al tornio da materiali tenaci). Esso è stabilito dall'equilibrio di due momenti che tendono a far ruotare l'ala in direzioni opposte. Il primo momento è dovuto, secondo l'osservatore a cavallo della samara, alla forza centrifuga agente sulle sezioni d'ala perpendicolari all'apertura. Ricordiamo sempre che la forza centrifuga è una forza apparente, vista solo da un osservatore che ruota. Un osservatore fermo vede invece la forza centripeta, che mantiene stabile il moto rotatorio. Consideriamo una striscia d'ala a metà apertura. La forza centrifuga è diretta all'infuori rispetto all'asse di rotazione della samara. Il momento di questa forza tende a far ruotare l'ala intorno al suo centro di massa in direzione tale da diminuire il diedro.

L'ala non assume una posizione orizzontale perché a ciò si oppone un altro momento, prodotto dalla forza aerodinamica, che tende invece ad aumentare il diedro. I due momenti si bilanciano in corrispondenza di un valore ottimale del diedro. Se una perturbazione casuale riesce ad alterare tale diedro, lo sbilanciamento dei momenti ripristina la situazione precedente.

In generale il diedro dovrebbe essere piccolo. Infatti se l'ala è quasi orizzontale il carico del disco si abbassa perché il diametro del cerchio spazzato dall'ala aumenta. Con ciò diminuisce la velocità di caduta. Ho trovato due samare d'acero che hanno approssimativamente il medesimo peso e le stesse dimensioni, ma assumono due diedri diversi. Quella con diedro minore scende più lentamente e quindi può venire trasportata più lontano dal vento, con vantaggio per la diffusione della specie.

L'ultimo problema di stabilità riguarda la direzione di caduta della samara. La maggior parte di queste scendono ruotando lungo un asse verticale. Alcune delle mie preziose samare si comportano diversamente. Anche quando la loro discesa non viene perturbata, percorrono una traiettoria elicoidale, sempre ruotando normalmente intorno al loro centro di massa. Norberg chiama «scivolamento laterale» questo movimento aggiuntivo. La forza aerodinamica agente su queste samare particolari fa inclinare il piano descritto dall'estremità alare mentre ruota intorno all'osservatore seduto sulla samara. Quando tale piano, normalmente orizzontale, si inclina, non viene riportato nella posizione iniziale, come accade con la maggior parte delle samare, bensì viene spostato lateralmente, in modo che il centro di massa della samara finisce per percorrere una traiettoria elicoidale rispetto al terreno. Il verso di rotazione seguito in tale spirale è opposto a quello seguito nella rotazione primaria. Se, guardando dall'alto, la samara ruota in senso orario, l'elica viene percorsa in senso antiorario. Il movimento elicoidale è molto più lento di quello di rotazione.

La figura in basso a destra nella pagina seguente riassume la mia analisi della distribuzione delle masse in una samara. Per trovare il centro di massa dell'intera samara l'ho messa in equilibrio su una lama, tenendo l'apertura perpendicolare al taglio e segnando poi una linea lungo di questo. Poi ho ripetuto l'operazione tenendo l'apertura parallela al filo della lama e segnando una seconda linea in corrispondenza di esso. L'intersezione delle due linee definisce la posizione del centro di massa.

È anche importante conoscere l'andamento della distribuzione della massa lungo singole strisce d'ala. Per questo ho tracciato su carta millimetrata la pianta dell'ala, dividendola poi in sezioni con la funzione di guida per tagliare a fettine la samara con una lama tagliente. Il baricentro di ognuna delle sezioni d'ala è stato poi determinato con la medesima procedura impiegata per l'ala intera e la sua posizione è stata riportata sul disegno dell'ala, in corrispondenza delle sezioni disegnate. Il risultato è riportato nella figura a fianco, in basso a destra. La distanza fra il centro di massa della sezione e il corrispondente bordo d'attacco è riportata nel disegno come percentuale della lunghezza della corda.

Angoli d'incidenza di un'ala che scende ruotando

Se la samara fosse un'ala piana, la distanza del centro di massa dal bordo d'attacco dovrebbe essere compresa fra il 27 e il 35 per cento della lunghezza della corda, per consentire una planata stabile. Nella samara che ho sezionato tale distanza era compresa fra il 26 e il 38 per cento. Un ottimo accordo con l'intervallo teorico, se si tiene conto degli inevitabili errori nella determinazione del centro di massa.

Per misurare la velocità di rotazione di una samara durante la discesa, ho usato una luce stroboscopica, illuminando un po' dall'alto il frutto che cadeva. Lavorando al buio, ho variato la frequenza dei lampi fino a che ognuno di essi mostrava la samara pressappoco nella medesima posizione. Un simile congelamento può venire ottenuto per diverse frequenze dei lampi, ivi compresa quella che corrisponde alla frequenza di rotazione della samara. Per esempio, una frequenza dei lampi pari alla metà di quella di rotazione consente di osservare il medesimo orientamento del frutto ogni due giri. Se la frequenza dei lampi è un terzo di quella di rotazione fa vedere la samara nella medesima posizione ogni tre giri.

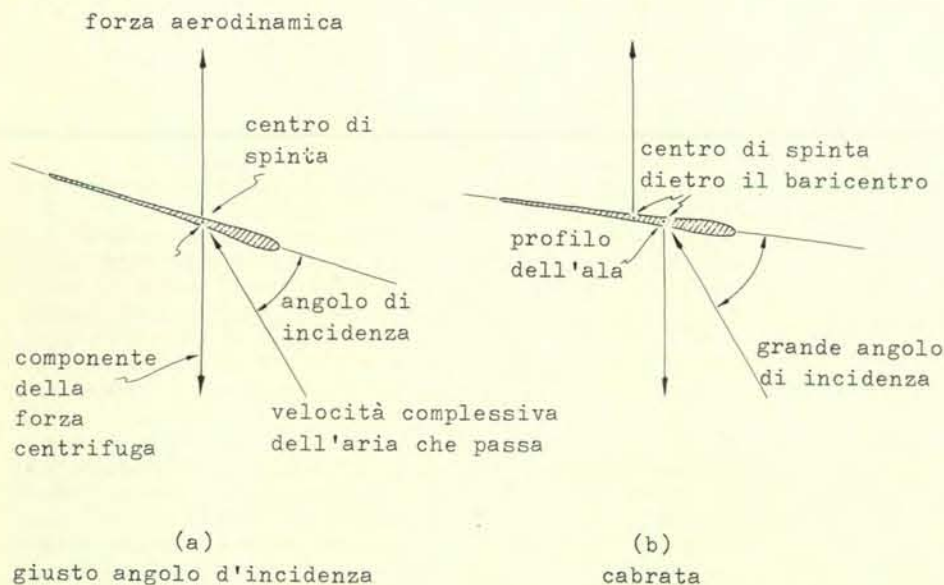
Per determinare la frequenza della rotazione si comincia con il lampeggiare a bassa frequenza, aumentando poi gradualmente la velocità del lampeggio. La frequenza più elevata che consente di fermare il movimento è quella uguale alla frequenza di rotazione. A frequenze più elevate si osserva costantemente la samara in posizioni diverse. Con questa procedura ho potuto osservare che la maggior parte delle samare che avevo raccolte aveva velocità di rotazione con frequenza compresa fra i 10 e i 12 hertz. Per fare un esempio, se si vuole osservare la samara in quattro posizioni diverse nel corso di una singola rotazione, si deve lampeggiare a una frequenza di circa 40 hertz.

L'illuminazione stroboscopica può venire impiegata per riprendere fotograficamente la discesa di una samara. In una stanza buia, i lampi successivi illuminano la samara in varie fasi della discesa, lasciando una registrazione permanente sulla pellicola di una macchina fotografica lasciata con l'otturatore aperto. McCutchen, invece, ha studiato il volo della samara di liriodendro (un parente della magnolia, originario del Nord America ma introdotto in Europa per scopi ornamentali) usando una luce continua proveniente dal basso.

La samara del liriodendro non si limita a ruotare intorno al centro di massa lungo la verticale di discesa. Rotola anche intorno all'apertura alare. Illuminando la samara dal basso nel corso della caduta, essa mostra alternativamente la faccia inferiore chiara e quella superiore scura.

La stabilità del volo di una samara

L'aggiustamento del diedro

La distribuzione della massa

Il fatto che dimostra come il frutto rotoli durante la discesa verso terra è documentato dalle immagini registrate sulla pellicola. Il contrasto può venire aumentato dipingendo di bianco una delle due facce della samara.

In questa rassegna ho parlato solo della samara di acero. Rimando al lavoro di McCutchen per la descrizione del volo di samare diverse. Sono particolarmente interessanti le samare di liriodendro e di frassino. È anche possibile sperimentare con modelli di samara in cartoncino, che consentono di osservare come variazioni nella forma e nella distribuzione dei pesi influenzino il volo.

Conservazione e restauro

Ricerche scientifiche multidisciplinari e materiali espressamente messi a punto consentono oggi di affrontare in termini nuovi il problema della conservazione e del restauro delle opere d'arte

di Paolo Parrini

Negli ultimi anni, con l'aumento dell'istruzione e con l'estendersi del turismo di massa, specialmente straniero, anche l'opinione pubblica si è dimostrata più sensibile, di quanto lo fosse in passato, ai problemi della conoscenza, della fruizione e della conservazione dei beni culturali. Questi ultimi, specialmente quelli costituiti da opere d'arte di facile comprensione e accessibilità, come quadri, statue, monumenti, sono ritenuti fonte di elevazione culturale e morale da più larghi strati della popolazione e questo interesse ha, oltretutto, una rilevanza economica di notevole entità sia nel settore pubblico sia in quello privato.

Si fanno sempre più frequenti le visite a musei, a mostre e a città particolarmente ricche di opere d'arte compiute o tramite la scuola o altre organizzazioni culturali e turistiche e si sta assistendo a un netto cambiamento nella considerazione popolare. Questo nuovo atteggiamento ha evidenziato maggiormente tutta una serie di problemi che intorno alle opere d'arte sono vivi da tempo e che oggi sono oggetto di attenzione da parte di gruppi di ricerca, di istituti, di singoli studiosi e delle istituzioni che sovrintendono al patrimonio artistico del paese.

A questa accresciuta sensibilità dell'opinione pubblica fa riscontro una maggiore dotazione di fondi da utilizzare negli interventi di conservazione e di restauro, la cui attuazione, grazie ai recenti sviluppi scientifici e tecnologici, va oggi considerata non più solo su basi di natura storica, artistica, filologica o artigianale, ma anche su basi scientifiche, con il ricorso ai mezzi più appropriati che le conoscenze attuali mettono a disposizione.

Da queste premesse trae motivo una vera e propria scienza di appoggio multidisciplinare ai problemi della conservazione e del restauro, la quale comporta approfondite conoscenze di chimica, fisica, metallurgia, biologia, informatica intese come un insieme di tecniche, di metodologie e di materiali espressamente messi a punto per risolvere i problemi dei beni artistici.

I vincoli del problema

Il problema della conservazione e del restauro delle opere d'arte è piuttosto complesso e difficile da risolvere, sia per l'oggettiva difficoltà della materia, sia per una serie di vincoli che impediscono allo scienziato o al restauratore di operare utilizzando tutte le possibilità offerte dalle conoscenze moderne.

Profonde differenze esistono, per esempio, nelle dimensioni del bene da conservare: si pensi al libro in pergamena miniata, all'oggetto di oreficeria, ai quadri di un museo o agli affreschi di una chiesa, ai grandi monumenti marmorei, o di altro materiale litoide, come chiese, chiostri, templi, archi di trionfo. Evidentemente tecniche valide in un caso diventano inapplicabili in un altro, quando le superfici e i volumi cambiano di migliaia di volte. Gli esami conoscitivi, così come gli altri interventi, assumono a volte dimensioni impensate e improponibili.

Il secondo vincolo è dovuto all'unicità del messaggio storico-artistico che l'opera d'arte incorpora e che impedisce di poter effettuare dalla stessa prelievi significativi, così che si è costretti a utilizzare o predisporre tutta una serie di tecniche microanalitiche, tali cioè da non comportare variazioni sull'opera dovute all'esame scientifico della stessa. Strettamente legati a questi sono i problemi della autenticazione e della datazione dell'opera, che oggi vanno fatte non solo con criteri estetici, ma anche e soprattutto scientifici.

Terzo vincolo, per coloro che progettano un restauro, è quello della reversibilità dell'intervento; esiste cioè la necessità di poter avere la possibilità di riportare l'opera d'arte nello stato in cui si trovava in qualunque momento lo si decida, senza cioè che l'intervento abbia modificato in modo permanente l'aspetto estetico e il messaggio artistico che l'opera contiene. È evidente che questo vincolo costituisce, per coloro che si occupano della messa a punto dei materiali o delle tecnologie, una delle maggiori difficoltà da superare. Significherebbe, per esempio, mettere a punto un adesivo che incolli perfettamente e permanentemente un antico incunabolo, vecchio di qualche decina di secoli e per il passare del tempo piuttosto mal ridotto, ma che possa essere tolto, se necessario, senza lasciar tracce sull'opera trattata. In genere gli adesivi attuali sono «buoni» proprio perché inalterabili nel tempo.

Quarto vincolo è quello del godimento pubblico del bene culturale e, per questo motivo, si è riluttanti a togliere l'opera d'arte dal contesto in cui si trova e per il quale l'artista l'ha creata. In alcuni casi l'oggetto viene tolto dall'esterno e posto in un museo, in atmosfera controllata, e in ambiente all'uopo studiato e predisposto; in altri casi ciò non avviene proprio perché in questo modo l'oggetto verrebbe sottratto alla fruizione pubblica. In altri casi ancora la sistemazione in un ambiente controllato è impossibile, o estremamente costosa, o con conseguenze estetiche non accettabili tanto che il ricovero conservativo non viene considerato. È il caso, per esempio, dei grandi monumenti di cui molte città e paesi d'Italia (da questo punto di vista una delle nazioni più ricche del mondo) vanno giustamente orgogliosi.

Questa serie di vincoli costringe, quin-

L'affresco di G. B. Ferretti *La moltiplicazione dei pani e dei pesci*, che si trova a Firenze, nella Santissima Annunziata, e di cui è riprodotto un particolare, presentava, prima del restauro *(a sinistra)* segni evidenti di degrado per la crescita di cristalli di solfato di calcio dovuta all'inquinamento. Il recupero dell'opera è stato totale e perfetto *(a destra)* ed è stato effettuato con un metodo messo a punto da E. Ferroni. L'intervento è consistito nella pulizia dell'affresco, nel successivo trattamento con carbonato d'ammonio e poi ancora con idrossido di bario.

In questa sequenza è riportata una fotografia al microscopio elettronico a scansione dei primi strati di un campione d'arenaria *(in alto)* e due X-grafie eseguite mediante microsonda a raggi X sullo stesso campione relative allo zolfo *(al centro)* e al calcio *(in basso)*. Lo zolfo è evidente già a una certa distanza dalla superficie, mentre l'assenza di zolfo e di calcio dalla stessa indica la scomparsa del solfato di calcio determinata dalla reazione dell'anidride solforosa presente nell'aria con il «cemento» calcitico intergranulare, prodottosi in seguito al dilavamento nei primi strati.

Questa fotografia illustra molto bene il degrado di un'arenaria lombarda: la pietra di Oggiono della Cappella Espiatoria di Monza. È evidente la formazione di scaglie superficiali a causa soprattutto di solubilizzazioni e di riprecipitazioni del «cemento» calcitico intergranulare dovute a percolazione di acqua piovana contenente sostanze aggressive come l'anidride solforosa.

TIPO DI ANALISI	DIMENSIONI DEL CAMPIONE	LIMITI DI RIVELABILITÀ (PARTI PER MILIONE)	PRECISIONE
SPETTROMETRIA DI EMISSIONE OTTICA	5-100 mg	$10\text{-}10^{3}$	5-10%
SPETTROMETRIA DI FLUORESCENZA A RAGGI X	100 mg	$10\text{-}10^{4}$	0,1%
ASSORBIMENTO ATOMICO	100-1000 mg	$10^{-3}\text{-}10$	0,5-2%
MICROSCOPIA ELETTRONICA	1-2 mm^2	100 000 INGRANDIMENTI	—
MICROANALISI ELETTRONICA A RAGGI X	1-2 mm^2	$10^{2}\text{-}10^{3}$	5%
ANALISI PER ATTIVAZIONE NEUTRONICA (NON DISTRUTTIVA)	50-1000 mg	$10^{-3}\text{-}10^{-1}$	2-15%
SPETTROMETRIA A PLASMA	100 mg	$10^{-4}\text{-}1$	0,5-2%
SPETTROMETRIA DI MASSA ORGANICA CON GASCROMATOGRAFIA	FRAZIONI DI MILLIGRAMMO	5×10^{-9} g	QUALITATIVA
SPETTROMETRIA DI MASSA INORGANICA	100 mg	$0,5\times10^{-3}\text{-}10^{1}$	5-20%
SPETTROMETRIA ELETTRONICA AUGER (AES)	1-2 mm^2	$10^{3}\text{-}10^{5}$	5-20%
SPETTROMETRIA FOTOELETTRONICA A RAGGI X (XPS)	1-2 mm^2	$10^{3}\text{-}10^{6}$	5-20%
SPETTROMETRIA MÖSSBAUER (SOLO PER Fe, Sn, Ru, Ni, Pt, Sb e Te)	QUANTITÀ CONTENENTE 20 mg DI ELEMENTO	$10^{2}\text{-}10^{3}$	SEMIQUANTITATIVA
GRAFIA MEDIANTE RAGGI X E GAMMA (NON DISTRUTTIVA)	—	—	—
SPESSOMETRIA E DIFFETTOSCOPIA A ULTRASUONI (NON DISTRUTTIVA)	—	—	—
TERMOVISIONE (NON DISTRUTTIVA)	—	—	0,2°C
OLOGRAFIA (NON DISTRUTTIVA)	—	—	—
FOTOGRAMMETRIA (NON DISTRUTTIVA)	—	—	—
FLUORESCENZA A RAGGI X INDOTTA DA RADIOISOTOPI (NON DISTRUTTIVA)	—	SEMIQUANTITATIVA	1-2%

Nella tabella sono riportate le principali tecniche utilizzabili attualmente per l'analisi dello stato di conservazione di un'opera d'arte e le limitazioni connesse a ciascuna di esse. Si tratta di tecniche usuali e di tecniche modernissime che richiedono costosi investimenti e alta specializzazione negli operatori. L'uso di tutte queste metodologie non è sempre possibile e non sempre è giustificato, però l'impostazione moderna della scienza di appoggio al restauro favorisce l'impiego di più tecniche per l'esame delle medesime grandezze al fine di ottenere ulteriori conferme nella individuazione delle stesse e di acquisire una visione plurilaterale dello stato di conservazione del campione sottoposto a esame.

di, a scegliere soluzioni tecnico-scientifiche appositamente individuate per la risoluzione di questi problemi e la consapevolezza dell'esistenza di questi vincoli sta comportando un vero e proprio salto di qualità negli interventi, sia in quelli di tipo conoscitivo, sia in quelli successivi (o che tali dovrebbero essere) di restauro o conservazione.

Il problema conoscitivo

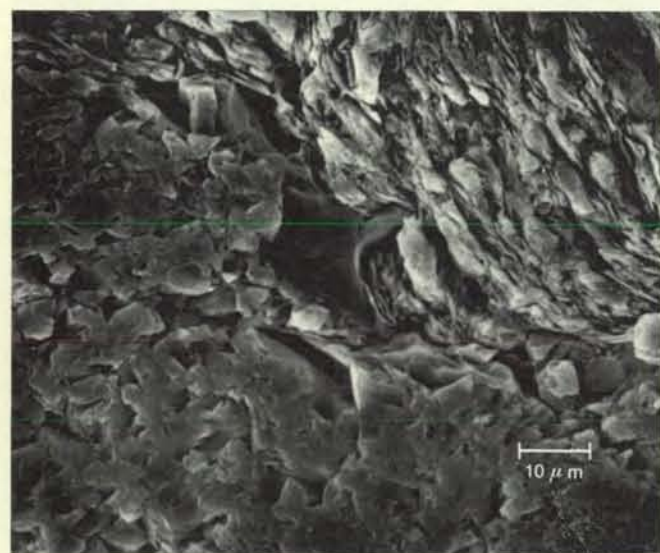
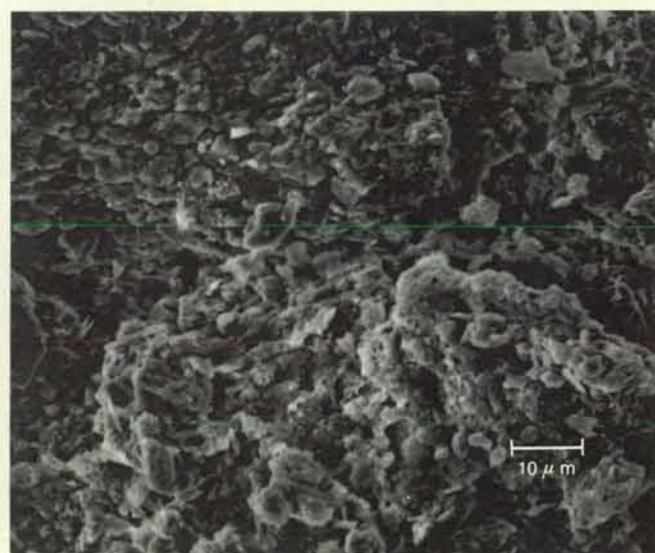
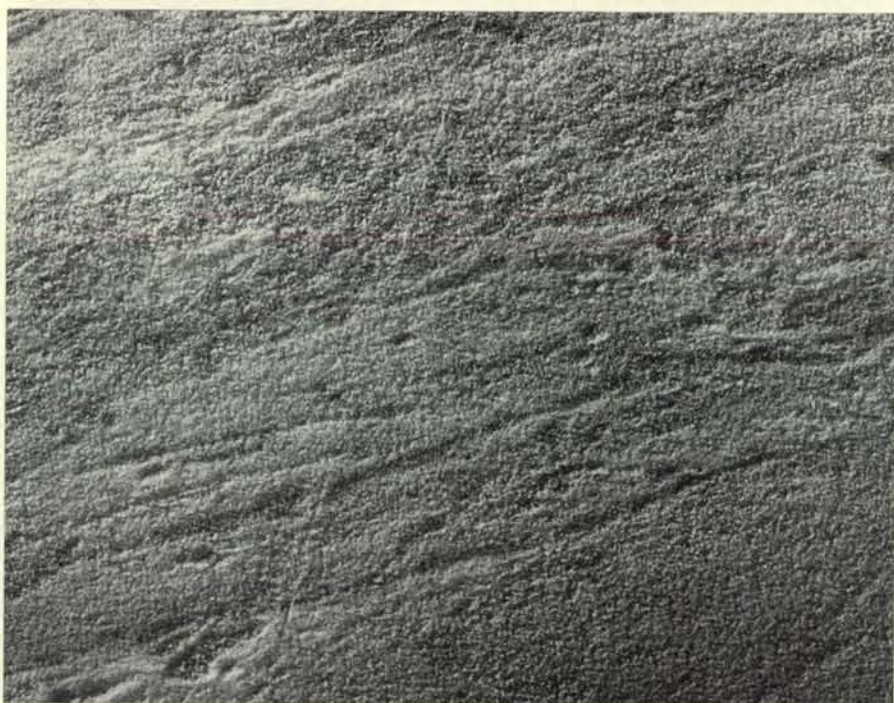
La base per ogni intervento sull'opera d'arte è naturalmente la conoscenza del suo stato fisico-chimico. Infatti solamente avendo ben presenti le condizioni in cui si trova dopo il trascorrere degli anni e dei secoli, si può pensare di porre in atto i metodi di conservazione che la tecnica moderna offre, oppure predisporre le azioni che portino al restauro, inteso come intervento che recuperi quanto più possibile l'«essenza» artistico-storica dell'opera.

A questo esame va aggiunto, secondo gli ultimi orientamenti della scienza dedicata ai beni culturali, lo studio accurato dell'ambiente, inteso come analisi del microclima e del microinquinamento, oggi considerati di estrema importanza nel ricostruire la storia del degrado già avvenuto, presente e futuro. Si intende come degrado tutto quell'insieme di fenomeni che hanno portato a un cambiamento nella struttura dell'opera d'arte, sia essa di pietra, di legno o di metallo. Si hanno così esempi di variazioni macroscopiche dell'aspetto superficiale delle pietre o dei colori di una tavola dipinta o dell'aspetto e del colore di un bronzo, oltre naturalmente alle variazioni microscopiche relative alla struttura. Questi effetti sono causati dal passare del tempo, dall'incuria, o se non altro dalla cattiva conservazione, a cui l'opera è stata sottoposta e, nel corso degli ultimi cinquant'anni, dalla presenza di sostanze inquinanti dovute sia alla forte industrializzazione del paese sia a cambiamenti drammatici di costume, come la diffusione del riscaldamento nelle abitazioni (prima a carbone poi a nafta e gasolio) e l'affermazione della circolazione automobilistica. A questi aspetti, trascurati negli anni recenti anche perché in contrasto con altri interessi, non viene neppure oggi data sufficiente considerazione se, per esempio, nella localizzazione delle centrali termoelettriche a carbone non si tiene conto che, per la natura del combustibile, si avrà inevitabilmente un aumento del tenore dell'anidride solforosa nell'aria malgrado gli apparati di disinquinamento. Tipico esempio di questo tipo di danno ambientale è quello che si sta verificando in India al famoso mausoleo del Taj Mahal per l'influenza di una raffineria costruita a soli 30 chilometri di distanza.

L'esame del microclima e del microinquinamento è oggi possibile con laboratori mobili automatici e con nuove apparecchiature che indicano agli specialisti le condizioni effettive del monumento, intese come condizioni potenzialmente nocive come potrebbe essere lo stato di «rugiada», cioè la situazione termoigrometrica che porta alla condensazione del vapor d'acqua sull'opera d'arte. È noto infatti da tempo, e altre analisi sono in corso, che le sostanze inquinanti come l'anidride solforosa (SO_2) o gli ossidi d'azoto (NO_x) si possono trasformare, in contatto con il vapor d'acqua, in acido solforico o nitrico (H_2SO_4 e HNO_3) i quali esercitano un'azione aggressiva sia su pietre e marmi, sia su bronzi o altri metalli. Nel primo caso si ha il cosiddetto fenomeno della «solfatazione», con la formazione di solfati di calcio, a scapito, per esempio, del carbonato di calcio ($CaCO_3$) che è un costituente di molte delle pietre utilizzate; nel secondo caso, si ha la formazione di nitrati o nitriti più o meno solubili. La concentrazione di questi inquinanti è alta all'inizio del processo di condensazione del vapore atmosferico così come all'inizio della pioggia, esercitando questi fenomeni atmosferici come una specie di azione di raccolta di tutte le sostanze accumulate e presenti nell'aria (con relativa dissoluzione nell'acqua), e diminuisce rapidamente nel tempo. Le sostanze inquinanti non sono poi più in grado di ritornare nell'aria quando, scomparse le speciali condizioni atmosferiche, si ha la fase di asciugamento o evaporazione, e così possono seguitare a compiere la loro azione, anche se lenta, sulle superfici su cui si sono inizialmente depositate. Questo processo non è esclusivo degli ambienti esterni, ma ha avuto e ha luogo anche in quelli interni. È famosa, per esempio, la formazione di cristalli di solfati di calcio sulla superficie degli affreschi del Beato Angelico in San Marco a Firenze. L'analisi e il riconoscimento del fenomeno ne hanno permesso, come vedremo, l'eliminazione e condotto al restauro integrale dell'opera.

Anche se non sempre è possibile un successo come quello indicato, l'analisi dello stato chimico-fisico dell'opera è garanzia di correttezza delle operazioni successive. Nel caso che sia possibile eseguire una sia pur minima campionatura, a livello di qualche centinaio di milligrammi, le metodologie analitiche moderne, ormai completamente strumentali, sono in grado di dare la composizione elementare fino a parti per milione e, in certi casi, fino ai fentogrammi (10^{-15} g). Diverse sono le tecniche utilizzabili: dall'assorbimento atomico alla spettrometria di massa, dalla spettrometria di emissione a quella mediante microsonda a raggi X e alla spettrometria gamma dopo attivazione neutronica. Alcune richiedono la distruzione del campione, altre invece non lo alterano macroscopicamente, limitandosi a registrare e interpretare cambiamenti di stato a livello atomico, limitati nel tempo e reversibili.

Le tecniche di analisi

L'assorbimento atomico conduce a un'analisi quantitativa che si basa sull'individuazione dei diversi elementi mediante lampade che contengono, ognuna, l'elemento che si presume sia presente nel campione. Si tratta di una analisi classica che permette però la determinazione di un solo elemento alla volta. È convenien-

Il confronto tra due immagini al microscopio elettronico a scansione di un campione di arenaria appartenente alla Cappella Espiatoria di Monza è significativo per comprendere i danni dovuti agli agenti atmosferici: a sinistra, il materiale integro ottenuto per taglio con lama diamantata a disco rotante; a destra, lo stesso tipo di materiale degradato dall'esposizione agli agenti atmosferici per circa 70 anni. È chiaramente visibile l'aumento dell'area esposta, dovuto allo scioglimento del «cemento» intergranulare che è stato attaccato dalla pioggia acida.

In questa fotografia di un campione di arenaria facente parte sempre della Cappella Espiatoria di Monza si nota in particolare il sollevamento della crosta superficiale e il colore virato verso il giallo-bruno, dal grigio-azzurro della pietra originale. Le analisi eseguite presso l'Istituto Donegani di Novara, con la spettrometria Mössbauer, hanno imputato il cambiamento del colore alla parziale formazione di ferro trivalente da ferro bivalente, sufficiente a originare il fenomeno.

Le due fotografie consentono di mettere a confronto l'aspetto superficiale di un marmo di Carrara, estratto dalla cava e sottoposto alla normale levigatura *(in alto)*, e dello stesso dopo circa 50 anni di esposizione all'aperto nella zona industriale di Novara. È evidente l'azione di degrado degli agenti atmosferici che hanno dilavato i prodotti di reazione della solfatazione (gesso).

te quindi un esame preliminare che può essere ottenuto con la spettrometria di emissione che consente di conoscere rapidamente, in modo semiquantitativo, tutti gli elementi presenti fino a livelli di qualche parte per milione.

Non deve sorprendere il fatto che si voglia spingere l'analisi fino a quantità così piccole che sembrerebbero trascurabili. Molte volte il riconoscimento di pezzi di opere d'arte o l'attribuzione di queste a «botteghe» o centri di produzione o a un determinato autore può avvenire con sicurezza proprio dal confronto delle tracce di elementi presenti come impurezze nei materiali usati.

La spettrometria di massa inorganica (così come quella organica se il campione è di questo tipo) si basa sulla capacità di un campo magnetico di deviare gli ioni di ogni elemento a seconda del rapporto fra la loro massa e la loro carica elettrica. In questo modo, una volta ionizzato il campione in esame, per esempio con una scintilla a radiofrequenza, si possono separare i vari elementi presenti e raccoglierli su una lastra fotografica, da cui si risale al tipo e alla quantità. Con questo strumento possono essere rilevate sostanze in concentrazione di 0,5 parti per milione.

La spettrometria a raggi X mediante microsonda utilizza invece un sottilissimo fascio elettronico che eccita i vari elementi presenti. Questi emettono radiazioni X che, analizzate nella loro frequenza caratteristica, permettono di ottenere in forma quantitativa l'analisi del campione.

Infine, l'analisi per attivazione neutronica si basa sui fenomeni di radioattività artificiale indotta negli elementi di un campione sottosposto a bombardamento di neutroni in un reattore nucleare. Una volta tolto dal reattore, il campione è radioattivo e, per spettrometria gamma, si risale al tipo e alla quantità di elementi presenti. Si riesce così ad avere la composizione elementare del campione con una precisione del ± 5 per cento con una notevole rapidità e più che altro utilizzando piccolissime quantità di sostanza (bastano 100 milligrammi).

Tenendo sempre presenti le dimensioni estremamente ridotte e la rarità del campione esaminabile, anche molti altri mezzi possono essere impiegati nello studio delle opere d'arte. Oltre a quelli della diffrazione e della fluorescenza dei raggi X, della spettroscopia molecolare nel visibile (ultravioletta e infrarossa), dei sistemi di cromatografia in fase liquida e gassosa, della spettroscopia a risonanza magnetica nucleare (NMR) e a risonanza paramagnetica elettronica (EPR), nonché di vari metodi microchimici condotti su microcampioni inglobati (utilizzati soprattutto per i dipinti), è di grande importanza un gruppo di tecniche, cosiddette «di superficie» che, molto recenti e ancora poco diffuse in Italia, permettono di dare la composizione elementare della superficie (almeno fino a qualche millesimo di millimetro in profondità) e, quello che più conta, lo stato di ossidazione in cui in essa si trovano gli elementi. In questo modo, non solo si può conoscere da quali elementi è costituita la superficie (che in un'opera d'arte molte volte è fondamentale), ma anche il suo stato chimico. Si può così conoscere lo stato di conservazione (se per esempio sono presenti ossidi, solfuri, solfati, prodotti di corrosione ecc.), che è di estrema importanza per la progettazione dei metodi di conservazione o, ancora di più, per gli interventi di restauro. Si tratta della spettroscopia elettronica Auger (conosciuta anche con la sigla AES) e della spettrometria fotoelettronica a raggi X (XPS) che, in genere abbinate, danno le informazioni che abbiamo sopra indicato. La spettroscopia AES non fornisce le informazioni complete sullo stato chimico in cui si trovano gli elementi presenti, ma la natura di questi, ed è basata sulla interazione di un fascio di elettroni a energia relativamente bassa (1-10 keV) con la superficie solida. Questo impatto può provocare l'emissione di un elettrone «Auger» che viene opportunamente rilevato e misurato. Con la spettrometria XPS, che invece rivela l'emissione di elettroni per irraggiamento della superficie con raggi X, si ottiene una serie di informazioni simili a quelle fornite dall'AES, ma in più si ha la determinazione dello stato chimico degli elementi, che ne costituisce il punto di forza. Un tipico spettro XPS è riportato nella illustrazione di pagina 63 in alto.

Un'ultima tecnica utilizzata per indagini su prodotti artistici (vetri, bronzi, ceramiche, vasi ecc.) è la spettrometria Mössbauer in cui si sfrutta l'effetto dell'assorbimento di raggi gamma per risonanza. Il nucleo atomico di un certo elemento agisce come sorgente radioattiva gamma quando è particolarmente eccitato, mentre assorbe questo tipo di radiazione nello stato fondamentale, come è il caso del campione durante l'analisi. L'applicazione analitica è limitata dal numero degli isotopi Mössbauer disponibili, in pratica solo quelli del ferro e dello stagno. Con questo metodo, però, si può conoscere tutto del ferro (stato di ossidazione, dimensioni delle particelle ecc.) in argille, ceramiche, pigmenti, dipinti e pietre. Il metodo è non distruttivo ed è applicabile all'esame delle superfici mediante la tecnica della riflessione. Queste informazioni possono essere utilizzate per il riconoscimento di campioni con caratteristiche simili (per esempio in studi di provenienza), oppure per chiarire le tecnologie di fabbricazione di certi vetri o ceramiche.

Quelle che abbiamo passato in rassegna sono quasi tutte le tecniche più sofisticate impiegate per l'esame della superficie o per l'analisi della composizione, utilizzanti dei campioni, sia pure estremamente piccoli (altrimenti si potrebbe, in teoria, utilizzare anche i vecchi metodi di analisi), che devono essere prelevati dall'opera d'arte in esame. Nessuna di esse infatti può operare su grandi dimensioni e, in genere, per una serie di ragioni, sono da considerarsi di tipo «distruttivo»,

In questa fotografia di un particolare di uno dei cavalli di San Marco in bronzo è possibile vedere come appariva prima della pulitura *(a sinistra)* e dopo il restauro *(a destra)*. È evidente il recupero dello splendore della doratura per eliminazione della patina del tempo e dei prodotti di corrosione superficiali. Sono nettamente visibili le «striature» che rendono famose queste statue.

In questa illustrazione è riportato l'andamento nel tempo di alcuni parametri caratteristici della pioggia. Il grafico in alto si riferisce all'andamento della pioggia nell'arco di 12 ore *(curva in colore intenso)*, con l'indicazione dell'intensità relativa della precipitazione *(curva in colore chiaro)*. Nei grafici in basso sono riportate le concentrazioni di alcune sostanze presenti nella pioggia. L'acidità libera viene data in equivalenti litro per 10^{-4} e la conducibilità elettrica in microsiemens per centimetro, mentre tutti gli altri parametri sono espressi in milligrammi per litro. Di particolare interesse il comportamento dell'inquinamento da anidride solforosa, la cui concentrazione è massima all'inizio della precipitazione e minima al termine. Il fenomeno del dilavamento dell'anidride solforosa presente nell'atmosfera avviene regolarmente solo se la pioggia cade senza interruzione. I grafici sono ripresi e rielaborati da un lavoro di F. Pantani, P. Cellini Legittimo, E. Recanelli e G. Sigg pubblicato sulla rivista «Inquinamento», 9 settembre 1977. I dati riportati si riferiscono a sette campioni di pioggia raccolti a Firenze il 13 novembre 1975 durante un fenomeno meteorico non continuo.

La sequenza fotografica riproduce un tracciato ricavato all'Istituto Donegani di Novara da uno spettrometro di massa inorganica. Dallo spettro ottenuto con sorgente a scintilla di un campione di bronzo sono rilevabili, oltre agli elementi base costituenti la lega (Cu, Zn, Sn, Sb), impurezze costituite da: Fe, As, Bi (100 ppm); Si, P (30 ppm); Na, Al, Zn, Ag, Pb, (10 ppm); Cd, Mn, Cr (3 ppm); V, Ti (0,1 ppm); Au (0,001).

anche se non in senso stretto dato che, dopo alcuni tipi di analisi, il campione può essere ulteriormente utilizzabile, se non è di ostacolo la sua piccola quantità.

Esistono però anche metodi conoscitivi detti «non distruttivi» i quali possono essere adoperati con successo, anche al limite sull'intera opera d'arte. Si tratta di metodi quali la radiografia a raggi gamma, la gammagrafia, l'esame a ultrasuoni, la termovisione, l'olografia interferometrica, la fotogrammetria, nonché la fluorescenza a raggi X indotta da radioisotopi utilizzate su statue, quadri, tavole, libri e altri oggetti d'arte.

La tecnica radiografica è comunemente nota: si tratta di irradiare, con apparecchi particolari che generano raggi «duri», l'opera da studiare. Celebre in proposito lo studio dei sarcofaghi egizi, racchiudenti le mummie sacre, «fotografati» con questo sistema prima di essere aperti in modo da operare in perfetta sicurezza.

Con i raggi gamma si sfrutta l'emissione di tali radiazioni da parte di quattro isotopi artificiali: cobalto 60, iridio 192, cesio 137, tulio 170. L'impiego di uno o di un altro isotopo è determinato dall'entità dello spessore e dal tipo di metallo; il vantaggio di queste radiazioni, rispetto ai raggi X, è che in genere l'operazione è più semplice e, dato che la sorgente può essere considerata come puntiforme ed emissiva su di un angolo solido molto più ampio di quello di una sorgente, della stessa potenza, a raggi X, si ha un effetto di appiattimento che può essere vantaggioso nel caso di opere a forti differenze di spessore.

Con gli ultrasuoni si possono determinare gli spessori di una statua o di un bassorilievo senza molta precisione, ma in compenso con molta rapidità e semplicità. Con speciali trasduttori che trasformano impulsi elettrici in onde meccaniche (di frequenza ultrasonica), inviando fasci di onde a frequenza ultrasonica sul campione, queste vengono riflesse dalle interfacce costituite da difetti interni o dalla stessa parete opposta del pezzo in esame. In tal modo, misurando il tempo di transito del fascio dal trasduttore alla parete opposta e ritorno, si può avere una misura abbastanza precisa dello spessore, la cui conoscenza può essere utile come fase preliminare di tutta una serie di interventi.

Per quanto riguarda la termovisione, mediante una speciale camera con lenti al germanio, si intercettano le radiazioni infrarosse presenti sull'oggetto inquadrato. In questo modo la sua immagine mostra, con una precisione di circa 0,2 gradi centigradi, le differenze di temperatura esistenti in quel momento. È possibile allora visualizzare zone dove in quell'istante esiste umidità di condensazione (zone fredde), minore velocità di evaporazione (zone meno fredde), dissesti nella statica degli intonaci (zone relativamente calde). Le esperienze effettuate (per esempio da Giovanni Urbani dell'Istituto centrale del restauro su alcune chiese) hanno dimostrato che questa tecnica ha prospettive di applicazione di grande interesse.

Sfruttando infine le note caratteristiche dell'olografia (si veda l'articolo *L'olografia nel campo del restauro* di Franco Gori e Giovanni Urbani in «Le Scienze», n. 74, ottobre 1974) è possibile ottenere un «ologramma» dell'opera d'arte in cui sono fissati in modo perfetto, data la luce monocromatica, tutti i particolari della superficie. Se gli ologrammi vengono eseguiti a tempi ben definiti, si possono evidenziare variazioni anche infinitesime dell'opera d'arte. Applicazioni in questo campo sono state eseguite da S. Amadesi, F. Gori, R. Gulla, P. Pasquini dell'Università dell'Aquila e G. Guattari dell'Università di Roma su tavole (per esempio il dipinto raffigurante Santa Caterina di Pier Francesco Fiorentino ad Assisi) dimostrando che è possibile evidenziare dei distacchi della parte pittorica dal supporto in modo assolutamente «non distruttivo».

A queste tecniche analitiche distruttive o non distruttive, che riassumiamo nella tabella di pagina 58, va aggiunto un mezzo attualmente fondamentale per l'ottenimento di risultati precisi e rapidi o per avere delle correlazioni che sarebbe praticamente impossibile avere con l'esame manuale. Intendiamo l'uso del calcolatore e dei programmi di calcolo, oggi essenziale sia nella guida delle varie apparecchiature, sia nella raccolta, sia nella manipolazione dei dati. Molti dei risultati odierni non sarebbero possibili senza l'uso di calcolatori sempre più sofisticati. Ma anche nell'interpretazione dei risultati l'uso di programmi particolari ha fatto compiere progressi prodigiosi alla scienza dei beni culturali.

Si considerino per esempio le tecniche di *pattern recognition* al calcolatore, utilizzate per elaborare la moltitudine di dati provenienti dalle analisi chimiche multielemento (analisi per spettrografia di emissione, in attivazione neutronica ecc.) nelle indagini sulla provenienza di manufatti ceramici, in vetro, in pietra o di metallo. Si tratta dello studio delle correlazioni e delle similitudini di insiemi di oggetti caratterizzati da molti parametri. Nel caso di oggetti caratterizzabili da due soli parametri, individuando ogni esemplare con due numeri in un normale diagramma cartesiano, gli oggetti simili tra di loro, rappresentati da un punto, si raggruppano in *cluster* o gruppi. Se ora sul diagramma si inseriscono i valori di un campione di cui non si conosce la provenienza, se questi si collocheranno in uno dei gruppi già presenti, esso può considerarsi simile agli altri e quindi della stessa provenienza. Se gli elementi in esame sono *n*, questo principio si può estendere a *n* dimensioni; in tal modo gli elementi di giudizio sono tanti e tali da poter raggiungere la quasi certezza dell'attribuzione. La risoluzione del problema in questo caso è affrontabile solo con un calcolatore, che provvede a confrontare ciascuno degli oggetti esaminati (rispetto a tutte le variabili esaminate) con tutti gli altri oggetti disponibili. Esso provvede anche a fare una scala delle similitudini e somiglianze e, quindi, a seconda del grado di similitudine, si avranno gruppi differenti. Un esempio di applicazione di questo metodo è quello adottato da E. Mello dell'Istituto Donegani di Novara relativamente alla discriminazione, in base al contenuto di alcuni elementi in tracce, tra le produzioni ceramiche medioevali di Savona, Albissola, Genova e Pisa, nonché all'attribuzione a questi gruppi di alcuni reperti di scavo non classificabili sulla base delle sole caratteristiche tipologiche.

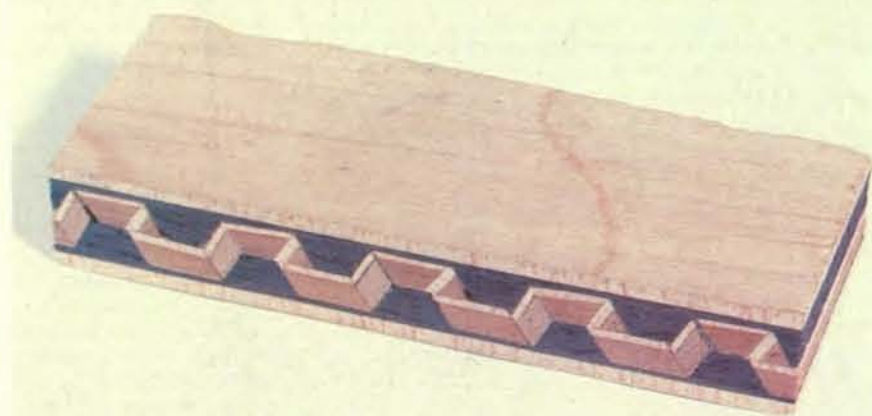
Il metodo è del tutto generale e può quindi essere utilizzato in molti casi concettualmente analoghi, non solamente in campo archeologico.

Spettri fotoelettronici a raggi X di due diverse porzioni, interna ed esterna, di un'incrostazione prelevata da un monumento lombardo in arenaria di Oggiono, la Cappella Espiatoria di Monza. Si noti la presenza di un picco dello zolfo nello spettro della porzione interna *(in alto)*. La sua posizione e la contemporanea maggiore intensità del picco del calcio suggeriscono che in questo campione sia presente gesso (solfato di calcio biidrato), prodotto di degradazione della pietra costituente il monumento.

La conservazione

Il compito principale di coloro che sono addetti alla tutela dei beni culturali è quello di preservare nel tempo per le generazioni future le opere d'arte a loro affidate. Una volta conosciuto lo stato in cui si trovano, si deve pensare a come mantenerle, in modo che non si deteriorino: compito questo reso sempre più difficile dagli aspetti della vita moderna da cui queste opere sono circondate e in cui, se si può dire, sono immerse. Il problema è diverso a seconda del tipo di opera da conservare.

Se si tratta, per esempio, di un gioiello o di una opera simile (smalti, sculture piccole, bronzetti), il compito è relativamente facile anche perché sono facilmente superabili tutti i vincoli di cui abbiamo parlato in precedenza. Le tecnologie disponibili attualmente permettono di mantenere in un ambiente chiuso, di volume relativamente piccolo, temperatura e umidità al valore voluto, così come l'intensità luminosa e la stessa atmosfera che può essere, per esempio, di un gas inerte come azoto o argo. Le cose si complicano quando si devono conservare opere più grandi e più delicate come, per esempio, una pala di altare o un grande dipinto. Molte volte è

Lo spettrometro Auger (AES) e fotoelettronico a raggi X (XPS) dell'Istituto Donegani di Novara è asservito a un calcolatore Univac 1100/20 per l'impostazione della sperimentazione, l'esecuzione delle misurazioni e l'analisi dei dati. Lo strumento è adatto all'esame delle superfici metalliche di cui è in grado di descrivere lo stato di corrosione sulla base dei prodotti che ne derivano.

Questo tipico pezzo di legno multicolore è stato utilizzato per il restauro della decorazione di un mobile toscano della fine del Quattrocento con il metodo «a toppo». L'unione di legni naturali di vario colore permette di ripristinare il disegno originale. Il legno nero è derivato da tronchi di quercia che hanno subito, nei secoli, un processo abbastanza spinto di carbonizzazione.

conveniente, dopo gli interventi di pulizia e di restauro, mantenerle dove sono sistemate da secoli, perché, se il loro degrado non è molto ampio, ciò significa che l'ambiente naturale, come potrebbe essere una chiesa, è quello che più si addice all'opera e questa può rimanere ancora per molto tempo nello stesso posto. Oltretutto, senza il fumo delle candele o delle lampade a olio che tanto hanno contribuito a «scurire» le opere pittoriche, la pittura dovrebbe risentire dell'ambiente esterno meno di prima (se si esclude l'inquinamento). Se l'opera è in museo, può essere messa in un ambiente condizionato e, almeno apparentemente, il problema è risolto. Apparentemente, perché negli ultimi tempi ci si è accorti che in una sala di museo, se l'opera d'arte è famosa, la presenza della folla di visitatori squilibra in modo netto il condizionamento creando condizioni di umidità pericolose sia per lo stato fisico, sia per quello biologico delle tele o del legno delle tavole. Non è infatti da trascurare l'attacco di muffe o batteri vari alle sostanze organiche che compongono il quadro o il libro famoso o la statua lignea. Recentemente si sono ipotizzate protezioni particolari a velo d'aria che, pur non impedendo il «godimento» dell'opera, la racchiudono come in una campana e costituiscono uno schermo invisibile, che impedisce in modo dinamico la penetrazione dell'aria inquinata, dell'umidità, ed esercita anche, qualora fosse necessario, un'azione disinfettante.

Se l'opera è di dimensioni ancora maggiori e non è possibile confinarla in un ambiente chiuso, come qualche volta è stato fatto, le tecnologie di conservazione presentano maggiori difficoltà e un minor grado di sicurezza. Esistono esempi di conservazione in ambienti particolari come quello di Stoccolma, in cui un intero edificio è stato costruito intorno alla nave ammiraglia della flotta di Gustavo Adolfo II, la *Wasa,* recuperata dal mare dove si inabissò durante il viaggio inaugurale il 10 agosto 1628. Essa viene continuamente spruzzata d'acqua di mare per impedire che il legno, con cui è completamente costruita, si asciughi e praticamente si polverizzi.

Ogni tipo di opera d'arte va protetto e conservato con metodi differenti. I metodi fisici sono quelli che più si prestano a superare molti dei vincoli di cui abbiamo parlato. L'esempio precedente si riferisce appunto a un metodo fisico, ma se ne possono citare tanti altri. Le statue in grande maggioranza vengono sistemate in musei e quindi in ambienti controllati. Di alcune si sono fatte copie da esporre nel luogo dove in precedenza era collocato l'originale (come è avvenuto per il *David* di Michelangelo a Firenze e per i cavalli di San Marco a Venezia); altre invece vengono messe direttamente in museo come avviene quasi sempre nel caso di reperti archeologici (per esempio, i due stupendi bronzi greci ritrovati sul fondo del mare prospiciente Riace in provincia di Reggio Calabria). Per quanto riguarda gli affreschi, in molti casi si è provveduto a «staccarli» e a rimontarli in museo. Se l'intervento è opportunamente eseguito, l'opera viene messa nelle condizioni ambientali ideali e può essere ammirata in tutta la sua bellezza. Vi sono opere però che restano esposte al godimento del pubblico e vedremo fra breve quali metodologie possono essere impiegate in questi casi.

Per le loro dimensioni o per il significato che la loro presenza comporta, queste opere restano dove da secoli sono state poste (si pensi alle grandi chiese, ai grandi monumenti in pietra, ai colonnati marmorei, vere e proprie storie illustrate) e quindi il problema della conservazione va affrontato in luogo e in tutta la sua drammaticità.

Un primo provvedimento è quello di eliminare quanto più è possibile le cause esterne di degrado come le vibrazioni provocate dal traffico automobilistico e il conseguente inquinamento da scarichi, deviando, per esempio, la circolazione; poi si deve procedere all'applicazione di sostanze protettive. Va detto in proposito che attualmente sono poche le sostanze con tutti i requisiti necessari ad assicurare a questo intervento una lunga durata. In genere si tratta di prodotti messi a punto per altri scopi e applicati, tra l'altro, alle opere d'arte. Solo ultimamente alcune industrie hanno iniziato ricerche su prodotti di sintesi specifici per questo impiego.

Mentre è relativamente semplice proteggere un quadro su tavola o su tela con particolari vernici, è molto difficile adottare una tecnica analoga per gli affreschi, per le pietre esterne e per le statue in bronzo. Molte volte le vernici protettive adottate non hanno avuto la stabilità nel tempo auspicata: si sono avuti casi di ingiallimento rapido o di accumulo di sporco a dimostrazione che la ricerca di prodotti nuovi appositamente studiati per questi usi con l'aiuto di analisi molto sofisticate è in questo momento una necessità molto sentita. Parlando di vernici protettive (a base di siliconi o di resine epossidiche) siamo passati nel campo degli interventi chimici. I più classici sono quelli a base di benzotriazolo per proteggere dalla corrosione le opere in bronzo o quelli che utilizzano strati protettivi sulla superficie sia di pietre sia di bronzi. Rispetto a questi ultimi, l'intervento può consistere nell'applicazione di resine sintetiche, i sostituti delle cere o delle sostanze grasse usate fin dalla antichità, che hanno però dei limiti nella loro deperibilità; si impiegano anche materiali polimerici sintetici per «proteggere», ad esempio, l'azione dell'inibitore di corrosione (come il benzotriazolo), dall'azione asportante degli agenti atmosferici.

Le tecniche chimiche di conservazione varieranno ovviamente a seconda del materiale di cui è costituita l'opera d'arte. Se si tratta di pietre, si cerca di evitarne il degrado rendendole quanto più è possibile idrorepellenti, perché è ormai accertato che l'acqua è il veicolo principale che permette l'attacco ai vari leganti dei materiali che costituiscono la pietra stessa (si veda in proposito l'articolo *La conservazione della pietra* di K. Lal Gauri in «Le Scienze», n. 120, agosto 1978). Il trattamento deve però consentire la naturale traspirazione della pietra e su questo punto molte sostanze usate negli ultimi tempi hanno condotto a risultati non soddisfacenti perché completamente impermeabilizzanti. L'uso di resine polimeriche ha comportato infatti la formazione di una pellicola che pone anche problemi estetici e di colore non indifferenti. Recentemente sembra che buone prospettive di successo si possano avere dall'uso di perfluoroeteri, appositamente sintetizzati dai laboratori della Montedison per questo tipo di applicazione.

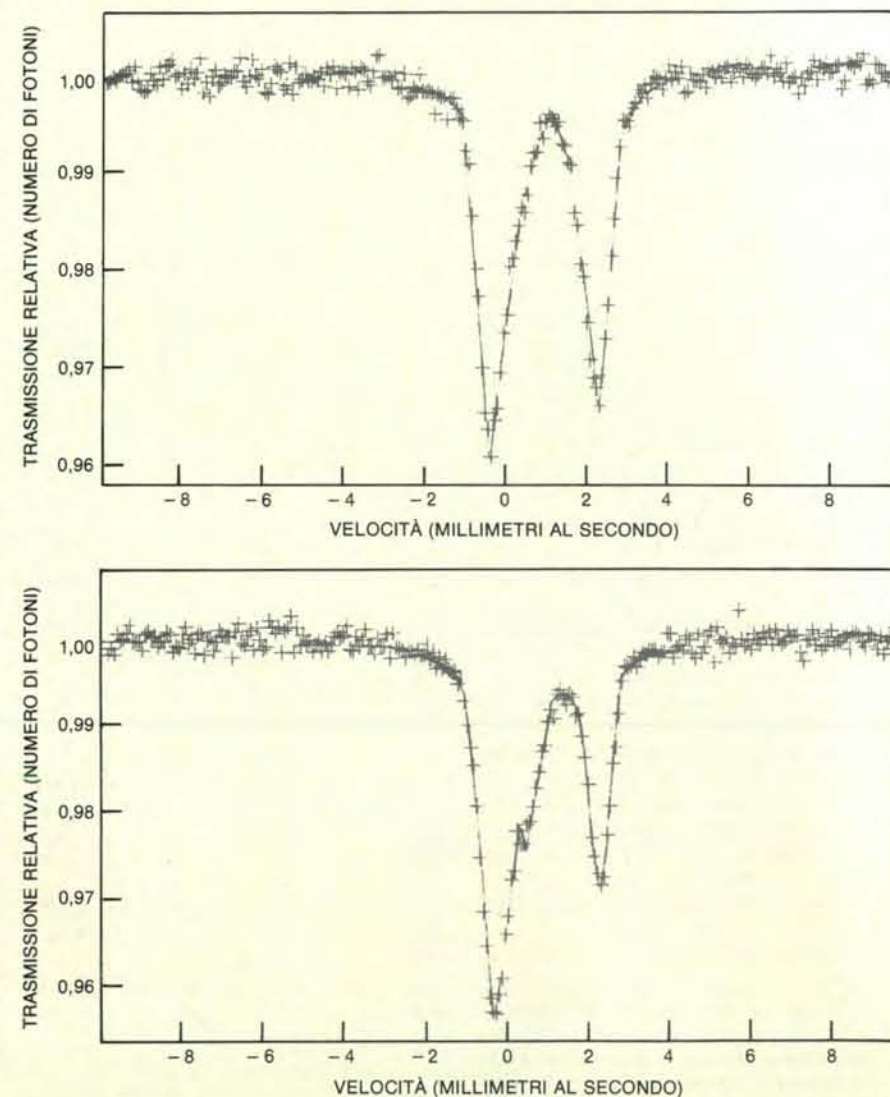
Se si tratta di bronzi, è fondamentale un'opera di pulitura, che non intacchi la «patina» formatasi nel tempo, con prodotti di passivazione del metallo e che in pratica lo protegge perfettamente. Quando questa patina dà segni di corrosione, imputabili in genere alle mutate condizioni ambientali, allora si interviene con inibitori di corrosione e vernici protettive finali.

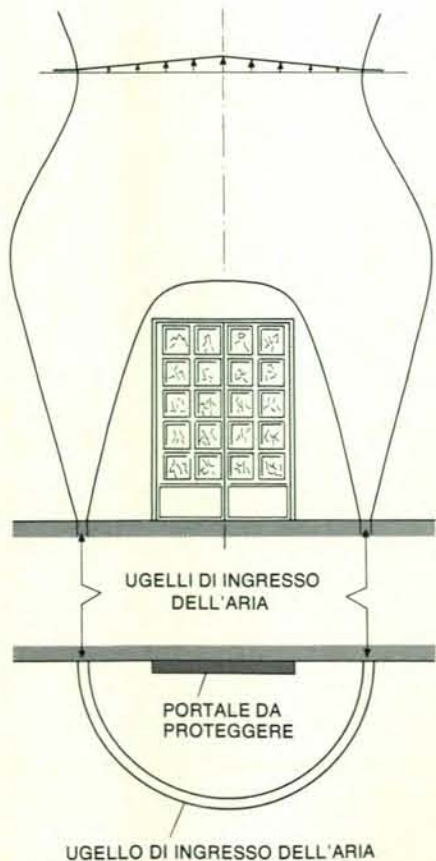
Per quanto riguarda la conservazione dei prodotti lignei (tavole dipinte e simili), si cerca di renderli meno igroscopici e quindi più stabili dimensionalmente alle variazioni di temperatura e di umidità dell'ambiente. Si sono usate a questo scopo impregnazioni con soluzioni di resine acriliche (Paraloid, Lucite ecc.), con nylon solubile (Caloton), nitrocellulosa, acetati di polivinile ecc., con effetti che però non sembrano giustificare in pieno un trattamento del genere, specialmente per quanto riguarda la stabilità dimensionale. Negli ultimi tempi si sono usate anche resine termoindurenti (poliesteri e resine epossidiche) che però hanno il difetto della totale irreversibilità. Tentativi recenti di impregnazione con particolari monomeri e successiva loro polimerizzazione con radiazioni ionizzanti sembrano aver avuto migliori risultati, anche se il procedimento non sembra molto agevole. Resta anche in questo caso la necessità di ricerche approfondite, da parte delle società chimiche, che consentano la messa a punto di prodotti specifici che, oltretutto, avrebbero non solo un successo scientifico, ma anche commerciale.

Per i manufatti lignei e per gli altri di origine organica, come arazzi, tappeti, tele, va considerata inoltre la minaccia costituita dai microrganismi, come le muffe, o dagli insetti, come tarli, termiti ecc. La lotta in questo caso è piuttosto difficile. Si sono adoperati antibiotici e disinfettanti vari a base di zinco, arsenico o di oli distillati dal catrame, ma va sempre tenuto presente che queste sostanze possono «sporcare» l'opera da conservare. Contro gli insetti xilofagi si è adottata la gassificazione con bromuro di metile o floruro di solforile, ma queste sostanze solo in pochi casi possono essere adoperate con sicurezza. Certi interventi poi sono quasi impossibili con i mezzi attuali: si pensi all'intervento contro i tarli che attacchino un coro ligneo o le architravi di una chiesa. Una linea di speranza viene dai nuovi studi sugli «insetticidi» non tradizionali, quali quelli basati sui «ferormoni» che, senza in alcun modo intervenire sull'insetto, fanno sì che esso venga deviato o catturato in modo da impedirne la riproduzione e quindi offrono delle possibilità finora impensate. Si tratta di una ricerca che ha portato a risultati eccezionali nei confronti di insetti nocivi all'agricoltura e alle foreste o alle derrate alimentari e che potrebbe finalmente offrire la possibilità di liberare aree anche vaste dagli insetti del legno.

Studi in questo senso sono in corso all'Istituto Donegani di Novara, uno dei pochi centri di ricerca al mondo specializzati in questo settore. Le difficoltà principali derivano dal dover individuare per ogni insetto il suo «attrattore» sessuale, se ne è provvisto, o cercare di conoscere con quale messaggio si incontrano specie di sesso diverso. Questo non è che il primo passo per arrivare alla sintesi del prodotto naturale con cui poter preparare trappole per mezzo delle quali catturare la quasi totalità di una specie degli insetti da eliminare. Se la ricerca avrà successo, interi musei potranno finalmente essere liberati dal pericolo (oggi più pressante a causa del riscaldamento invernale) costituito dagli xilofagi.

Il restauro

Quando un'opera d'arte ha subito l'azione del tempo ha bisogno in maggiore o minor grado di interventi di ripristino comunemente indicati come «restauro». Il problema del restauro è piuttosto complesso perché implica non solo soluzioni tecniche, ma anche artistiche. Non è più pensabile, oggi, allo stato attuale delle conoscenze, che il restauro riporti l'opera alle condizioni originarie, perché questo costituirebbe, in un certo senso, un falso, dato che l'opera come si vede attualmente non può essere disgiunta dalla «patina» del tempo, cioè da quella serie di modifiche cromatiche ed estetiche che il passare del tempo impone a tutte le cose. Conoscendo però la storia della sua creazione e possibilmente la metodologia adottata dall'artista, il restauro deve arrivare a riportare l'opera al suo stato «artistico», così che sia salvaguardato lo spirito che porta racchiuso nella sua bellezza. Esiste uno stretto legame fra arte, storia e scienza che deve essere rispettato per evitare che si intervenga in modo irrimediabile, modificando per sempre l'opera artistica. Il legame fra arte e storia è piuttosto comprensibile. Purtroppo è influenzato dal gusto dell'epoca, specialmente nei tempi passati in cui mancava la critica scientifica d'arte e si tentava il restauro tenendo pre-

Questi spettri Mössbauer relativi al ferro sono stati eseguiti alla temperatura dell'azoto liquido presso l'Istituto Donegani di Novara e si riferiscono a due zone di pietra arenaria provenienti dalla Cappella Espiatoria di Monza: una interna *(in alto)* e una superficiale *(in basso)*. La presenza in quest'ultima di un picco relativo al ferro trivalente è un tipico esempio di azione di degrado.

A fianco della fotografia della formella di Lorenzo Ghiberti, il *Sacrificio di Isacco*, realizzata in occasione del Concorso del 1401 per le porte del Battistero di Firenze *(a sinistra)* è riportata la gammagrafia eseguita presso l'Istituto Donegani di Novara su incarico dei Laboratori dell'Opificio delle pietre dure di Firenze. La gammagrafia ha permesso di accertare che nella formella non esistono saldature e che l'arca *(in basso a destra nelle fotografie)* è stata eseguita dall'artista con la tecnica dell'incastro e a bulino, lavorazione tipica degli orafi.

Nel disegno è illustrato un sistema brevettato di protezione a velo d'aria di un portale, visto frontalmente *(in alto)* e in pianta *(in basso)*. Dall'ugello semicircolare esce un getto d'aria particolarmente sagomato e a velocità calcolata tale da mantenere l'opera come racchiusa in una campana, respingendo gli elementi aggressivi esterni. Se l'aria è depurata e in condizioni controllate si evita quasi completamente il contatto con gas nocivi come l'anidride solforosa.

sente il gusto del momento e non quello che aveva avuto l'autore dell'opera. Si è assistito così a restauri compiuti nei tempi passati che hanno completamente sconvolto e modificato l'opera originale e oggi si è costretti a correggere e a eliminare le conseguenze di tali operazioni.

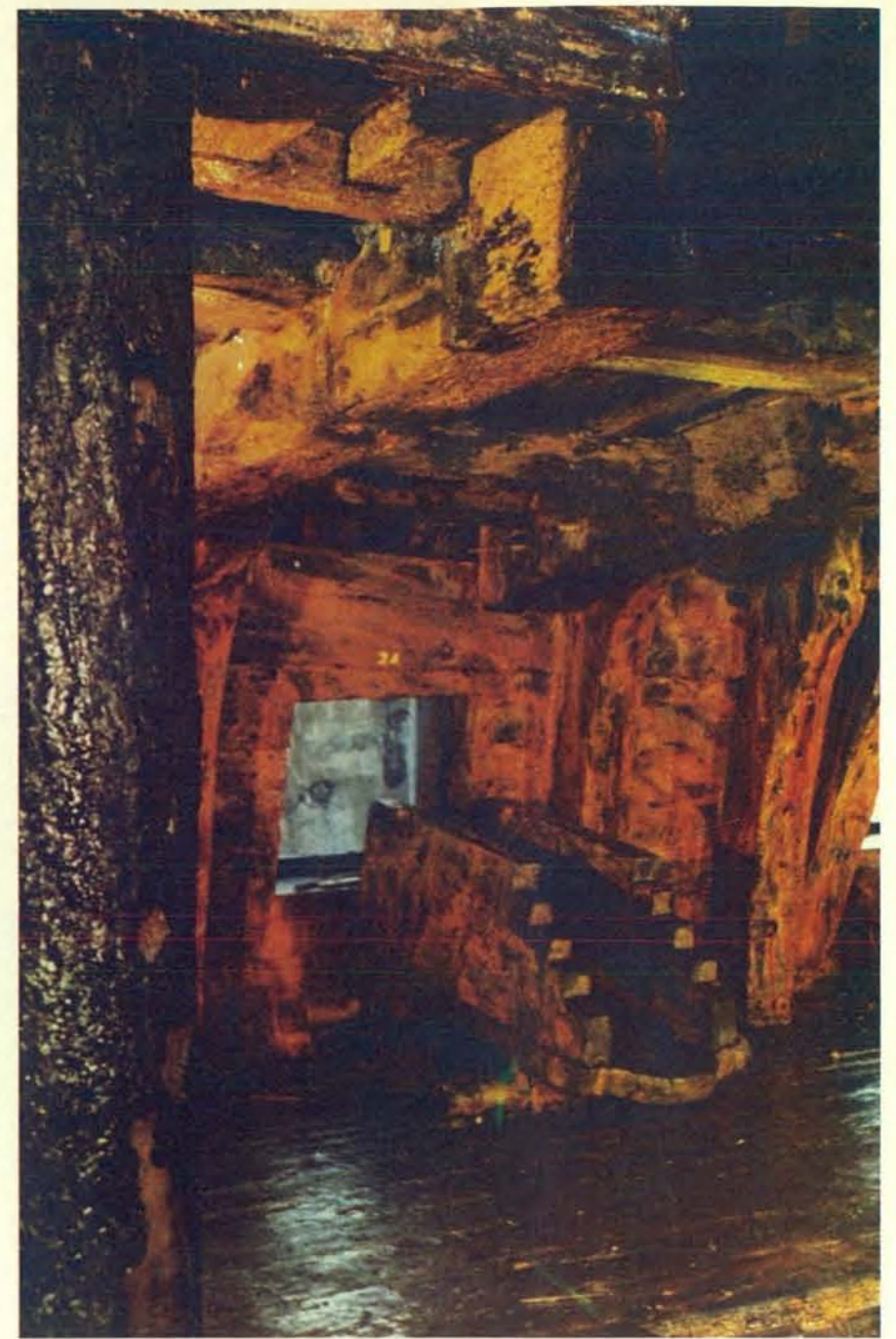
Per questi motivi l'azione di restauro va compiuta con grande senso di responsabilità verso la storia e, specialmente, con le conoscenze scientifiche che oggi l'operatore è in grado di avere. Il legame storia-arte va perciò integrato con quello scientifico, che è il solo che permette al restauratore di operare in condizioni quanto più possibile obiettive e tali che possano essere modificate in futuro, in caso di nuove scoperte storiche o di nuovi materiali che siano più stabili o in qualche aspetto migliori di quelli attuali.

L'esame dell'oggetto da restaurare rimane perciò la base di tutto il lavoro successivo e questo per due ragioni fondamentali. La prima, di tipo puramente tecnico: è necessario conoscere il tipo di materiale utilizzato, il suo stato di conservazione e le eventuali modifiche subite nel corso del tempo. La seconda, di tipo tecnico-storico: non si può prescindere dalle circostanze in cui l'opera d'arte è nata, dalla provenienza dei materiali originali e, in particolar modo, dalla tecnologia impiegata.

La conoscenza dei materiali permette ovviamente di utilizzare quelli più adatti a un corretto intervento così come la conoscenza operativa e tecnica di un manufatto ne può permettere la corretta manutenzione come è il caso, ad esempio, del restauro delle decorazioni lignee di un mobile quattrocentesco. Poiché, però, non si chiede nell'intervento di restauro, soprattutto di tipo pittorico, alcuna azione che possa essere imitativa o falsificante, è inutile pensare di fare uso o di ricercare materiali antichi che magari hanno dimostrato labile conservazione nel tempo o non adeguata stabilità.

Attualmente, molto si sta facendo per il ripristino delle opere d'arte utilizzando nuovi materiali la cui richiesta è sempre più pressante. Nel caso delle pietre e del marmo, le maggiori necessità sono sentite per materiali consolidanti che blocchino il degrado in atto e permettano di riportare l'opera alla sua essenza artistica. In genere si interviene sui fenomeni di perdita di coesione, seguiti da fratture abbastanza nette, con resine sintetiche termoindurenti dotate di buona resistenza all'invecchiamento; ma un materiale che, oltre a consolidare le varie particelle fra di loro, si comporti come la pietra sia dal punto di vista delle variazioni dimensionali, per effetto della temperatura e dell'umidità, sia degli scambi con l'ambiente come avviene normalmente, è ancora oggetto di ricerca. Per i fenomeni di perdita diffusa di coesione si usano in genere metodi chimici che sfruttano reazioni che tendono a ricostituire nella pietra degradata un «cemento» di natura silicea o a base di carbonati, simile a quello originario e oggi scomparso per l'alterazione dovuta in genere a solfatazione del carbonato di calcio. Trattamenti a base di silicati alcalini od organici, di fluorosilicati, di barite ecc. sono fra quelli più noti. Dal punto di vista scientifico, il ritorno allo stato chimico-fisico iniziale è certamente la strada più corretta, anche se, molte volte, non permette il ripristino di vuoti, fratture o di altri danneggiamenti.

Un esempio da manuale di questa metodologia di intervento si è avuto nel restauro degli affreschi del Beato Angelico in San Marco a Firenze e del Ferretti nella chiesa della Santissima Annunziata sempre a Firenze. In questo caso la solfatazione dello strato di intonaco, con formazione di gesso, era così avanzata da formare una serie di cristalli di solfato di calcio che, emergendo dallo strato pittorico, deturpava completamente i dipinti. La soluzione del problema è consistita nel ripristino del carbonato di calcio trasformatosi in gesso. Gli studi effettuati da E. Ferroni dell'Istituto di chimica fisica dell'Università di Firenze hanno portato all'individuazione di una serie di reazioni di «ritorno», che hanno permesso non solo la formazione del carbonato di calcio originario, ma addirittura la formazione di un carbonato molto più stabile all'attacco dell'anidride solforosa, dando una garanzia che il fenomeno di degradazione non si sarebbe più riprodotto nel tempo. Con tecniche sperimentali molto sofisticate, i restauratori dell'Opificio delle pietre dure di Firenze, diretti dal sovrintendente U. Baldini, hanno trasformato il gesso in solfato di ammonio mediante reazione con carbonato d'ammonio. Una volta completata la reazione, un trattamento con idrossido di bario ha trasformato il carbonato di calcio, così ricomposto, in carbonato di bario, dando un marcato consolidamento a tutta la superficie, senza con questo annullare la porosità dell'affresco che permette un regolare scambio con l'ambiente, garanzia di durata nel tempo, oltre a consentire la possibilità di eventuali futuri interventi. Non sempre però possono essere adottate soluzioni di questo genere e con un successo così completo.

Nel caso di dipinti su tavola, si può arrivare persino a eliminare il supporto quando non è più in grado di fungere come tale e lo strato pittorico, così separato, si riapplica su di una tavola nuova o anche su di un supporto di materiale diverso che deve avere le caratteristiche del supporto originale senza averne i difetti di comportamento nel tempo; per le tele si adotta la tecnica del rifodero, incollando una nuova tela sul retro del dipinto; per gli affreschi, quando non è possibile un intervento sul posto, è ormai di uso comune il «distacco» dal muro e la successiva applicazione a un supporto che dia le garanzie di conservabilità ormai perdute da quello originale. Durante queste operazioni l'introduzione di nuovi materiali ha permesso ai restauratori di intervenire in modo più sicuro e più «reversibile» che nel passato. Citiamo per esempio l'uso di lastre di resine poliesteri rinforzate con fibre di vetro come supporto degli strati pittorici, come è avvenuto per il *Crocifisso* di Cimabue, conservato in Santa Croce, dopo l'alluvione di Firenze; il ricorso a carte speciali, a frattura programmata, che permettono l'eventuale futuro distacco del dipinto dalla tavola di legno senza sottoporlo più al trauma del primo intervento o a paste sintetiche poliolefiniche, come quelle prodotte dalla Montedison, che, sostituendo quelle cellulosiche nei trattamenti chimico-fisici descritti in precedenza, danno una maggiore garanzia di durata, di assoluta imputrescibilità e di resistenza alle muffe e batteri.

La *Wasa*, di cui si vede l'interno, è l'ammiraglia di re Gustavo Adolfo II. Recuperata dal mare nel 1961, viene da allora mantenuta sotto continua pioggia di acqua di mare in un edificio costruito appositamente e situato nel porto di Stoccolma, allo scopo di evitare il suo deterioramento. Nella fotografia è chiaramente visibile un affusto di cannone, in legno, del ponte inferiore.

Lunga sarebbe la descrizione delle tecniche adottate dai vari centri di restauro. Certamente la tendenza attuale è quella di utilizzare metodi che conservino senza compromettere troppo, che lascino quanto più è possibile l'opera invariata, che operino in modo differenziato in modo da eliminare il dubbio della falsificazione. Per i materiali ci si orienta sempre di più verso i polimeri sintetici ai quali si chiede stabilità nel tempo che ancora, nei casi favorevoli, nessuno ha potuto sperimentare se non con metodi induttivi di tipo accelerato.

Numerosi sono ancora i problemi non risolti: certo che se all'impegno dei restauratori e alla loro sensibilità artistica si unirà sempre di più lo sforzo della scienza nell'analisi dello stato di conservazione e nella preparazione di materiali appositi per i vari interventi, il problema del restauro di opere, nate dalla sensibilità degli uomini e dal gusto del bello, sarà vissuto come un normale intervento e non, come spesso succede, come un'operazione difficile, incerta, in qualche modo sacrilega.

Gli schemi di colorazione delle ali delle farfalle

I disegni delle ali sono formati da un mosaico di minuscole squame che si dispongono secondo più di 100 000 modelli il cui sviluppo è però regolato da poche e semplici leggi che sono comuni a tutte le specie

di H. Frederik Nijhout

Poche cose nella natura possono uguagliare in bellezza e varietà i disegni delle ali delle farfalle. Questo ordine di insetti - i lepidotteri - comprende circa 100 000 specie ognuna delle quali si distingue dalle altre esclusivamente per i disegni e i colori delle ali. Il fenomeno diviene ancora più notevole quando si esamina il modo in cui tali disegni si formano, come io ho fatto per anni, e si scopre che tutto sommato la risposta è piuttosto semplice.

I biologi si sono domandati a lungo quali sono le basi fisiologiche della disposizione altamente differenziata di colori e disegni sulle ali (che chiameremo da ora in poi schema di colorazione) poiché esemplificano un problema fondamentale di morfologia animale. Quali processi regolano lo sviluppo di una ben determinata disposizione spaziale dei tessuti, degli organi e delle appendici che caratterizzano una data specie animale e all'interno di questi processi quali sono le differenze essenziali che danno a ogni specie il suo aspetto caratteristico? C'è ragione di credere che lo sviluppo degli schemi di colorazione si ottenga, per lo meno in linea di principio, attraverso processi del tutto simili a quelli che guidano lo sviluppo dei caratteri morfologici, poiché tutto lo sviluppo è fondamentalmente il risultato dei progressivi cambiamenti nell'espressione dei geni. Negli schemi di colorazione questa espressione ha come risultato la sintesi localizzata di molecole relativamente semplici di pigmento.

Le squame delle ali dei lepidotteri (nome che significa ali squamose) sono molto piccole e formano uno strato sulla membrana alare. Ogni squama è l'escrescenza appiattita di una sola cellula e misura circa 100 micrometri di lunghezza e 50 micrometri di larghezza (un micrometro è un millesimo di millimetro). Le squame si sovrappongono come tegole e coprono completamente la superficie membranosa dell'ala con una densità che va da 200 a 600 squame per millimetro quadrato. Chiunque abbia tenuto una farfalla per le ali sa che le squame si staccano facilmente e che i disegni spariscono con esse; si scopre allora che la membrana alare è incolore e traslucida.

Esaminando un'ala a forte ingrandimento diviene evidente che il suo disegno è in realtà un finissimo mosaico, in cui ogni tessera è una squama e ogni squama è di un solo colore. In genere i disegni sono composti da pochi colori, di regola tre, quattro o cinque. Piccole gradazioni di colore o di tono si possono poi ottenere per mezzo di variazioni nelle proporzioni tra i differenti tipi di squame, esattamente come le fotografie che illustrano questo articolo vengono riprodotte per mezzo di un gran numero di punti di quattro diversi colori. Per esempio, il disegno delle ali anteriori della farfalla *Precis coenia* è composto da squame di soli quattro colori: nero, giallo-arancio, rosso o marrone. Le squame bianche al centro della grande macchia a forma di occhio sono prive di colore; il loro biancore è dovuto alla riflessione della luce da parte di piccolissime bolle d'aria contenute all'interno delle squame. Benché ogni squama sembri avere un solo pigmento, la quantità di pigmento può variare; in *Precis coenia* si possono distinguere al microscopio squame di almeno tre intensità differenti di marrone. Analogamente, sembra che il pigmento delle squame giallo-arancio e rosse sia lo stesso con la differenza che nelle prime è più diluito che nelle seconde.

I colori delle squame delle farfalle possono essere dovuti alla presenza di pigmenti o a modifiche strutturali. Nel primo caso i pigmenti coinvolti sono pochi, prevalentemente melanine. Tutti i neri e i grigi e la maggior parte dei colori rossicci, bruni, rosso-brunicci e gialli sono forme di melanine. In ordine di abbondanza, la seconda categoria di pigmenti alari è quella delle pterine, che danno luogo a molti tra i rossi brillanti, gli arancioni, i gialli e alcuni bianchi. In qualche specie sono stati trovati anche pigmenti non identificati, che sono stati chiamati flavoni, carotenoidi e ommocromi; i colori di cui sono responsabili vanno dal gialliccio al bruniccio.

Alcuni tra i colori più brillanti non sono dovuti a pigmenti chimici, ma a piccole caratteristiche strutturali delle squame. Ho già nominato le macchie bianche a forma di occhio; oltre a queste, tutti i colori iridescenti o metallici, tutti i blu e i verdi delle farfalle diurne (ma non i verdi delle farfalle notturne) sono strutturali. Questi colori si originano per interferenza, diffrazione o diffusione della luce.

Le squame iridescenti devono il loro colore soprattutto all'interferenza della luce riflessa. A seconda della specie, i colori di interferenza si originano quando la luce viene riflessa da solchi orizzontali posti su numerose creste verticali presenti sulla superficie della squama o da file orizzontali di sottili lamelle che compongono la cuticola esterna della squama. Le squame iridescenti portano anche un pigmento, in genere una melanina scura, che assorbe molta della luce trasmessa e fa in modo che la componente riflessa appaia particolarmente brillante.

Si è visto che i diversi schemi di colorazione delle ali dei lepidotteri seguono una serie di regole piuttosto lineari. Queste regole sono state enunciate negli anni venti in modo indipendente e quasi con-

Nella microfotografia della pagina a fronte sono riprodotte le squame delle ali di una grossa farfalla notturna, *Hyalophora cecropia*, ingrandite 150 volte. La zona fotografata è vicina alla macchia oculare che si trova presso il bordo esterno delle ali anteriori dell'insetto. Le squame bianche sono in realtà prive di colore e il loro biancore è dovuto alla riflessione della luce da parte di minuscole bolle d'aria che si trovano all'interno delle squame. Le squame che vanno dal marrone rossiccio fino al giallo arancio contengono differenti quantità dello stesso pigmento.

La serie di disegni in scala crescente della pagina a fronte raffigura in dettaglio la struttura delle squame di *Morpho rhetenor*, una farfalla dell'America meridionale. I disegni sulla sinistra rappresentano il lato superiore dell'ala che nel maschio si presenta di colore blu metallico. Il colore è strutturale, deriva cioè dall'interazione della luce con le squame delle ali, ed è dovuto alla struttura delle creste presenti sulla loro superficie. Queste hanno numerosi solchi orizzontali le cui superfici superiori e inferiori distano circa 0,22 micrometri: circa la metà della lunghezza d'onda della luce blu, il che fa in modo che questo tipo di lunghezze d'onda interferisca costruttivamente quando la luce viene riflessa dalle loro superfici. I disegni a destra mostrano il lato inferiore dell'ala della farfalla, che è marrone; questo colore dipende dai pigmenti contenuti nella squama.

temporaneamente da B. N. Schwanwitsch dell'Università di Pietrogrado (ora Leningrado) e Fritz Süffert dell'Università di Friburgo. Essi scoprirono che moltissimi schemi di colorazione si possono far derivare da un solo archetipo o modello base. Questo viene generalmente chiamato il modello base dei ninfalidi, poiché sia Schwanwitsch che Süffert l'hanno scoperto per la prima volta in alcune specie di ninfalidi, una vasta famiglia di farfalle diurne.

Con gli anni è apparso chiaro che da questo modello base si possono fare derivare la maggior parte degli schemi di colorazione delle ali delle più importanti famiglie di farfalle diurne e notturne. È però necessario mettere in evidenza che il modello base non deve essere visto come uno schema ancestrale di colorazione, ma come uno schema puramente ipotetico da cui, per mezzo di opportune manipolazioni, si possono ottenere gli schemi delle specie reali.

La caratteristica saliente del modello base dei ninfalidi è una fila di ocelli (piccole macchie a forma di occhi) che si trova lungo il bordo dell'ala. Essi sono composti da una serie di cerchi concentrici il cui centro è sempre situato sulla linea mediana di una cellula alare (il termine cellula alare si riferisce a una porzione della membrana alare delimitata da venature ed è ben differente dalle cellule delle squame). Il modello base ha un ocello in ognuna delle cellule alari periferiche.

Nella regione centrale dell'ala vi sono numerose bande colorate che corrono dal margine anteriore dell'ala a quello posteriore, chiamato sistema centrale di simmetria. Esso consiste in due serie di bande di pigmento poste a destra e a sinistra di una linea immaginaria che corre all'incirca attraverso la metà dell'ala. La sequenza di colori che si osserva su una metà dell'ala è l'immagine speculare della sequenza sull'altro lato. I colori della banda sono quindi simmetrici rispetto alla linea mediana dell'ala.

Generalmente sulla linea mediana del sistema di simmetria centrale si trova una cospicua banda pigmentata, la macchia discale, la cui posizione serve normalmente per definire l'asse di simmetria. In molte specie di farfalle notturne la macchia discale si sviluppa in una grande macchia circolare a forma di occhio. Infine, uno o più sistemi di simmetria si trovano alla base dell'ala. Questi consistono in serie accoppiate di bande colorate come il sistema centrale, anche se i sistemi basali sono molto meno elaborati.

Non è nota alcuna specie che mostri l'intero modello base. Esso è infatti lo schema ideale da cui gli schemi reali possono venire fatti derivare mediante l'espressione o l'esagerazione di alcuni elementi e la soppressione di altri. Gli ocelli marginali sono di regola presenti solo su alcune cellule alari e la banda prossimale del sistema centrale di simmetria (la banda più vicina al corpo) è spesso mancante. Inoltre le bande del sistema centrale di simmetria possono essere molto irregolari, e non c'è alcuna relazione evidente tra quelle distali (le più lontane dal corpo) e quelle prossimali. Ogni elemento dello schema si può esprimere con un'ampia varietà di pigmenti o di colori strutturali per produrre forti contrasti oppure sottili gradazioni di colore.

La modificazione più notevole del modello base implica alcuni spostamenti laterali di determinati segmenti del sistema centrale di simmetria. Schwanwitsch ha chiamato questo fenomeno dislocazione, perché l'effetto visivo è sorprendentemente simile alla dislocazione in geologia, dove alcuni strati di rocce sedimentarie appaiono spostati lungo una frattura. Nei lepidotteri la dislocazione descrive l'interruzione di una banda di pigmento nel punto in cui questo attraversa una venatura e lo spostamento laterale di una notevole parte della banda lungo la venatura stessa. Schemi di dislocazione specie-specifici sono responsabili della maggior parte delle variazioni negli schemi alari. In numerosi esempi la dislocazione ha come effetto quello di generare uno spiegamento discontinuo di chiazze e macchie colorate. La dislocazione ha anche luogo sugli ocelli che quindi non sono sempre perfettamente allineati.

Il modello base dei ninfalidi, un modello teorico da cui si possono far derivare molti degli schemi di colorazione delle ali delle farfalle diurne e notturne, è stato elaborato indipendentemente negli anni venti da B. N. Schwanwitsch dell'Università di Pietrogrado e da Fritz Süffert dell'Università di Friburgo. I due modelli vengono qui rappresentati in modo semplificato, quello di Schwanwitsch a sinistra, quello di Süffert a destra, con l'indicazione delle affinità e delle differenze. Da tali modelli base si possono far derivare gli schemi di colorazione di circa il 75 per cento delle farfalle.

È molto probabile che parecchi aspetti del «progetto e sviluppo» degli schemi di colorazione dei lepidotteri possano essere spiegati mediante i correnti concetti della biologia dello sviluppo. Durante gli ultimi dieci anni molti biologi dello sviluppo hanno finito per accettare la proposta di Lewis Wolpert della Middlesex Hospital Medical School di Londra secondo cui, per poter proseguire correttamente lo sviluppo, le cellule devono giungere ad avere informazioni su dove si trovano (si veda l'articolo *Lo sviluppo delle forme negli embrioni* di Lewis Wolpert in «Le Scienze», n. 124, dicembre 1978). Deve dunque esistere un meccanismo per mezzo del quale le cellule si rendono conto della loro posizione in un organo o in un tessuto in via di sviluppo, utilizzando in seguito queste informazioni posizionali per iniziare o interrompere qualunque processo biochimico sia specificato dai loro geni per cellule in quella posizione. Per lo sviluppo degli schemi di colorazione questo significherebbe che una cellula di squama dovrebbe conosce-

In questa fotografia si possono osservare dei derivati del modello base. *Stichophthalma camadeva*, una farfalla diurna *(in alto a sinistra)*, è la più vicina al modello base. *Saturnia carpini*, una grossa farfalla notturna *(in alto a destra)*, ha un sistema centrale di simmetria evidente e una macchia discale ben sviluppata, ma solo un accenno di ocelli marginali.

***Cressonia juglandis*, uno sfingide *(in basso a sinistra)*, ha uno schema più ridotto, tipico di numerose farfalle notturne, con un sistema centrale di simmetria. *Taenaris rothschildia*, una farfalla diurna tropicale *(in basso a destra)*, ha uno schema molto ridotto, composto da ocelli marginali che si sono sviluppati in grandi strutture sproporzionate.**

re la propria posizione su un'ala per produrre il pigmento adatto.

Sviluppando la sua teoria dell'informazione posizionale, Wolpert ha ipotizzato che, dal momento che le cellule possono solo sapere dove si trovano rispetto a qualcos'altro, qualsiasi sistema in via di sviluppo deve avere almeno un punto di riferimento. Egli non ha specificato la natura biochimica o fisiologica di questi punti e buona parte degli sforzi teorici si è sviluppata per definire i tipi di processi e fenomeni che potrebbero svolgere una simile funzione. Quasi tutti i modelli che sono stati sviluppati ipotizzano l'esistenza di un gradiente di qualche tipo: un fattore fisico o una sostanza chimica che muta gradualmente di dimensioni o di concentrazione all'interno di un campo di sviluppo. I modelli più semplici da visualizzare presuppongono l'esistenza di fonti distinte di una sostanza chimica diffusibile: un morfogeno. La concentrazione di morfogeno sarebbe più elevata nelle cellule intorno e vicino alla fonte e gradualmente più bassa nelle cellule progressivamente più lontane. A seconda della concentrazione, il morfogeno potrebbe influenzare in maniera diversa l'espressione di determinati geni nella cellula attraverso cui diffonde.

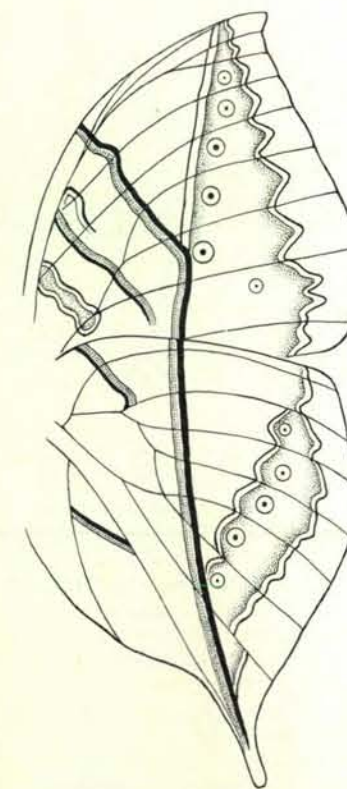
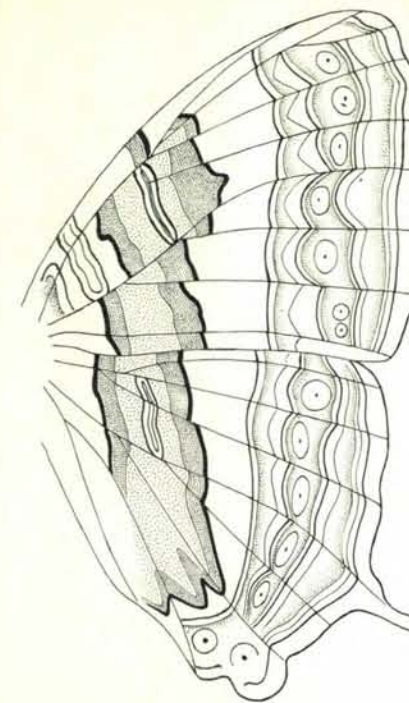
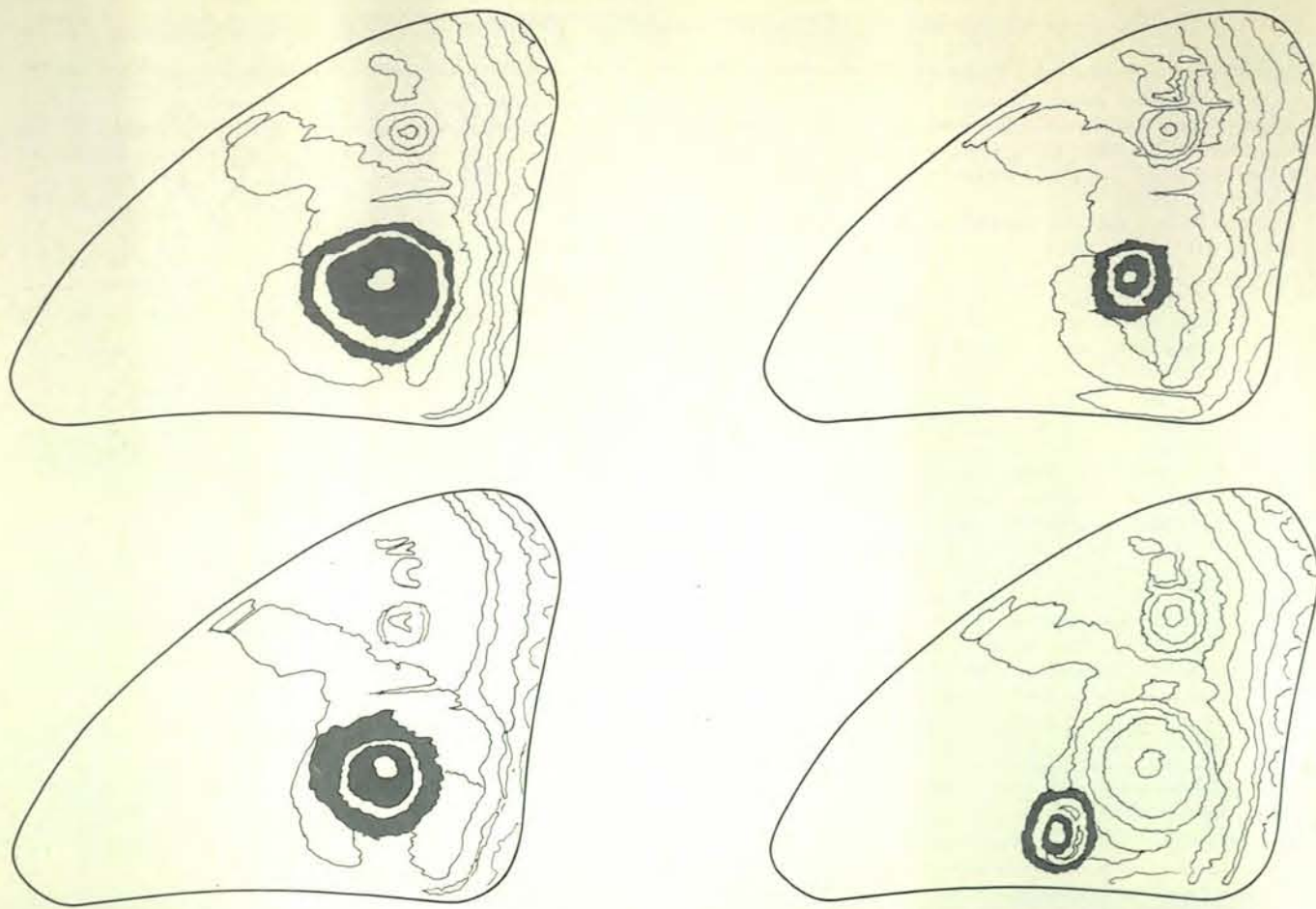
Accettando per il momento quest'idea, supponiamo che i geni interessati codifichino alcuni enzimi che entrano in gioco nelle vie di biosintesi di vari pigmenti. I disegni più semplici che si possono ottenere in queste condizioni sono anelli e cerchi. Tutto ciò che è necessario è che le cellule si trovino sulla fonte di informazione posizionale o entro una certa distanza critica da essa per differenziarsi allo stesso modo. In questo caso tali cellule sintetizzerebbero lo stesso pigmento.

Ho già messo in evidenza come i cerchi e gli anelli di pigmento siano caratteristiche comuni degli schemi di colorazione dei lepidotteri, sotto forma di ocelli marginali e di macchie oculari discali. Forse si potrebbe quindi utilizzare l'idea dell'informazione posizionale come modello per iniziare l'analisi dello sviluppo delle macchie oculari. In questo modello lo sviluppo di macchie oculari circolari dipenderebbe in qualche modo da un'attività specializzata delle cellule poste al centro.

Dovrebbe essere evidente che i processi fisiologici che determinano la posizione e la forma degli schemi di colorazione devono essere operanti mentre l'ala non è ancora pigmentata. Dunque uno dei problemi pratici di maggiore importanza quando si conducono lavori sperimentali sulla determinazione degli schemi di colorazione è la capacità di prevedere con esattezza il punto dell'ala su cui un elemento dello schema si svilupperà. Per alcuni anni ho lavorato sullo schema di colorazione di *Precis coenia*. Questa specie è particolarmente utile per tali analisi poiché ha vari punti di riferimento sulla copertura delle ali della pupa, lo stadio di sviluppo precedente all'adulto. Con l'aiuto di tali contrassegni è possibile prevedere la posizione di varie parti dello schema di colorazione presunto delle ali sottostanti non pigmentate molto prima che tali schemi si sviluppino veramente.

La posizione presunta del centro della grande macchia oculare sull'ala anteriore è particolarmente facile da identificare. È quindi possibile penetrare attraverso la cuticola pupale con un sottile ago da cauterizzazione e distruggere il gruppetto di cellule al centro della macchia oculare. Se l'operazione viene compiuta in un momento abbastanza precoce dello sviluppo (alcune ore dopo l'impupamento), quando l'ala si è appena sviluppata, lo sviluppo della macchia oculare è inibito in maniera specifica e totale. La stessa operazione eseguita 48 ore o più dopo l'impupamento non produce alcun effetto. Distruggendo le cellule in periodi diversi, ma non inferiori a 48 ore, si sviluppano macchie oculari di dimensioni intermedie. L'ala non si pigmenta fino al quinto giorno dello stadio pupale (circa 24 ore prima dell'apparizione della farfalla adulta); quindi la determinazione dei disegni precede la loro differenziazione (la sintesi dei pigmenti) di circa tre giorni.

Se lo sviluppo delle macchie oculari dipende da qualche proprietà particolare delle cellule poste al centro della macchia, dovrebbe essere possibile trasferire questa proprietà in punti diversi dell'ala per mezzo di un trapianto di cellule. Un simile trasferimento è stato provato con successo: quando ho trapiantato una piccola parte di epidermide dall'area che avrebbe formato il centro di una macchia oculare in un diverso punto dell'ala (da cui era stata tolta una porzione di epidermide di dimensioni simili), ho osservato che intorno a essa si è sviluppato un disegno pigmentato di forma circolare. L'epidermide trapiantata da altre regioni dell'ala non ha dato origine a un disegno di questo tipo. Sembra dunque che la determinazione delle macchie oculari sia una proprietà intrinseca di un piccolo numero di cellule (forse di una sola cellula) e che esse siano una delle fonti di informazione posizionale postulate da Wolpert, almeno per un elemento dello schema di colorazione. Ho chiamato fuoco un tale gruppo di cellule.

Ho già accennato che le ali anteriori della farfalla *Precis coenia* hanno solo quattro pigmenti, tutti melanine. Anche se non si conoscono le differenze molecolari tra queste sostanze che appaiono di colore diverso, è probabile che la sintesi di ciascuna di esse avvenga a partire da un precursore semplice rappresentato da un solo tipo di enzima, una tirosinasi. Sembra anche che tirosinasi con specificità differente possano servire per ciascuno dei quattro colori della melanina. Se fosse così, l'azione di un fuoco sarebbe, almeno in linea di principio, piuttosto semplice. Il fuoco farebbe in modo che le cellule sotto la sua influenza passino, per esempio, dal-

Gli esempi qui fotografati indicano di quanto sia possibile allontanarsi dal modello base. Gran parte dei cambiamenti derivano dalla dislocazione, cioè dall'interruzione di una banda di pigmento nel punto in cui questa incrocia una venatura alare e dallo spostamento laterale della banda lungo la vena. Sono qui rappresentati due casi, ciascuno esemplificato mediante la fotografia di un'ala e un disegno di accompagnamento. La farfalla *Cyrestis cocles*, raffigurata in alto, mostra la dislocazione delle bande del sistema centrale di simmetria in modo chiaro, ma poco sviluppato. La farfalla *Kallima inachus*, raffigurata in basso, ha evoluto un disegno che la fa assomigliare a una foglia morta sul suolo della foresta, imitazione perfetta delle venature fogliari e delle macchie di funghi.

Due esperimenti indicano le modalità di controllo dello sviluppo delle macchie oculari in una farfalla. Gli esperimenti sono qui illustrati con la farfalla *Precis coenia*. Lo schema normale della farfalla *(in alto a sinistra)* cambia notevolmente se si distrugge mediante cauterizzazione un piccolo gruppo di cellule poco dopo l'impupamento: quattro ore dopo l'impupamento, le macchie non si sviluppano *(in alto a destra)*; la cauterizzazione effettuata 24 ore più tardi ha come risultato lo sviluppo di una macchia oculare di dimensioni medie *(in basso a sinistra)*; se la cauterizzazione viene invece eseguita oltre 48 ore dopo l'impupamento non provoca alcun effetto. Se le cellule centrali non vengono distrutte ma trapiantate in un altro punto dell'ala, si sviluppa un disegno circolare *(in basso a destra)*. Questo non è però una vera macchia oculare, poiché occorre del tempo affinché il tessuto si cicatrizzi nell'area di trapianto e pertanto resta meno del periodo di 48 ore necessario per indurre lo sviluppo di una macchia di dimensioni complete.

la produzione di una tirosinasi per la melanina gialla a quella di una tirosinasi per la melanina nera. È probabile che tutte le macchie oculari siano organizzate intorno a fuochi di questo tipo e che il sistema di ocelli marginali si sviluppi intorno a una fila di fuochi. Nel modello base dei ninfalidi ci sarebbe dunque un fuoco per ogni cellula alare e ogni fuoco sarebbe situato sulla linea mediana della cellula.

È probabile che il sistema centrale di simmetria si generi anch'esso intorno a una fila di fuochi. In molte specie questo sistema consiste in una serie di cerchi concentrici fusi tra di loro; i centri dei cerchi, come quelli degli ocelli marginali, si trovano sulla linea mediana delle cellule alari. Inoltre ci sono prove sperimentali che dimostrano l'esistenza di fuochi funzionali nel sistema centrale di simmetria. Negli anni trenta Alfred Kühn e Melitta von Engelhardt dell'Università di Gottinga hanno compiuto una serie di esperimenti in cui hanno studiato l'effetto della cauterizzazione in vari punti del sistema centrale di simmetria della tignola grigia della farina *(Ephestia kühniella)*. Analizzando i difetti nello schema di colorazione che si sviluppavano dopo l'intervento essi dimostrarono che il sistema centrale di simmetria di questa farfalla viene generato da almeno due, ma probabilmente tre, «fonti di determinazione», di cui una si trova vicino al margine anteriore dell'ala, una vicino al margine posteriore e forse una vicino al centro dell'ala. Esperimenti analoghi svolti da altri studiosi su specie di farfalle diverse hanno confermato questi risultati.

Come il sistema centrale di simmetria, anche i sistemi basali di simmetria sono spesso formati da serie di cerchi concentrici fissi tra loro, ciascuno dei quali ha il centro situato sulla linea mediana di una cellula alare. Sembra dunque che il modello base dei ninfalidi si possa interpretare nel modo migliore come organizzato intorno a tre file di fuochi che corrono dal margine anteriore dell'ala a quello posteriore. Una fila distale di fuochi controlla lo sviluppo degli ocelli marginali, una fila centrale il sistema centrale di simmetria e una fila prossimale il sistema basale di simmetria.

Le bande di un sistema di simmetria si possono considerare originate dalla fusione degli anelli periferici di grandi ocelli. Un inconveniente di questa ipotesi è che le bande del sistema centrale di simmetria sono più frequentemente linee diritte piuttosto che archi di cerchio. Anche gli ocelli marginali spesso non sono circolari.

Prima di iniziare a descrivere lo sviluppo di disegni non circolari intorno a un fuoco, riassumerò le leggi che, secondo i miei studi, regolano lo sviluppo dei disegni sulle ali dei lepidotteri. A quanto si sa, i disegni sulle superfici superiore e inferiore dell'ala si sviluppano in modo indipendente. Gli elementi dello schema si sviluppano intorno a un numero finito di fuochi. Nella maggior parte dei casi un fuoco determina solo il disegno della cellula alare su cui è situato, perciò lo schema di colorazione di ciascuna cellula alare si determina indipendentemente da quelli delle altre cellule alari. Lo schema di colorazione non risulta quindi solo da un mosaico di squame, ma anche da un mosaico di cellule alari che rappresentano unità di sviluppo indipendenti. È possibile che un fuoco determini lo schema di varie cellule alari adiacenti, ma solo se in esse mancano fuochi attivi.

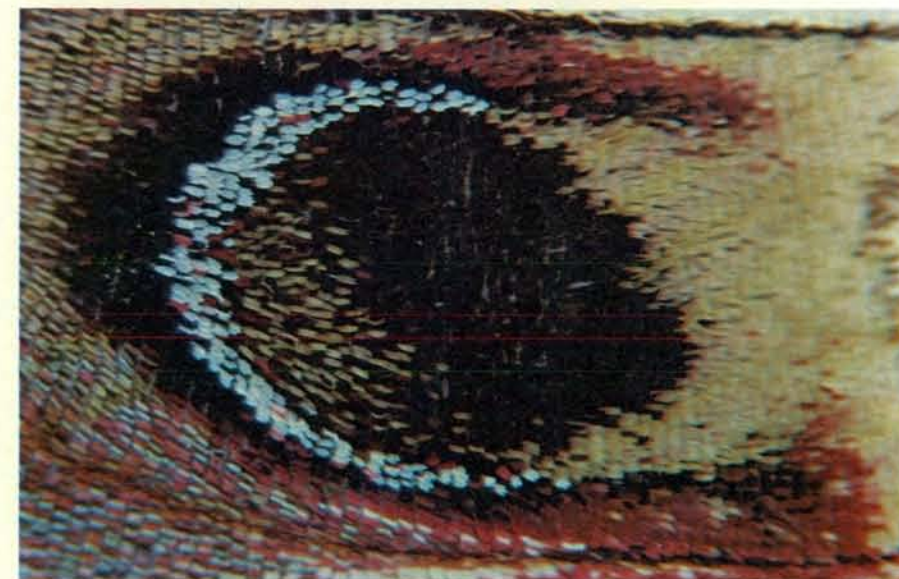
Schemi specie-specifici all'interno del piano di base dei ninfalidi si originano, almeno in parte, per variazioni nel numero di fuochi realmente espressi e nella loro posizione sull'ala. Gli spostamenti laterali dei fuochi e i disegni che essi generano in ogni cellula alare spiegano il fenomeno della dislocazione. Lo sviluppo di schemi specie-specifici dipende anche da cambiamenti nelle regole che governano la forma dei disegni che si sviluppano intorno a un dato fuoco.

Anche se sulle ali dei lepidotteri possono comparire dei cerchi, questi costituiscono ancora una minoranza tra i disegni alari; gli ocelli marginali assumono spesso la forma di ellissi, di mezzaluna, o forme più complesse. Tali strutture si osservano anche nelle bande del sistema centrale di simmetria. Un aspetto interessante dei disegni non circolari è che sono quasi sempre simmetrici rispetto all'asse di simmetria che corre a metà della cellula alare, parallelamente alle venature alari e attraverso un fuoco. I disegni non circolari sono spesso limitati a una sola cellula alare, come i normali ocelli circolari.

Per lo sviluppo di disegni circolari, secondo la teoria dell'informazione posizionale, sono necessarie due condizioni: in primo luogo il segnale posizionale si deve trasmettere con uguale velocità e intensità in tutte le direzioni a partire dal fuoco, in secondo luogo tutte le cellule delle ali devono seguire le stesse regole quando interpretano l'intensità del segnale posizionale. È chiaro che almeno una di queste condizioni deve venire violata quando intorno a un fuoco si sviluppa un disegno non circolare.

Nella sua forma più pura, la teoria di Wolpert dell'informazione posizionale presuppone che il segnale posizionale non muti; a sostegno di questa ipotesi esistono numerose osservazioni e prove sperimentali. L'idea è che possa esistere un solo segnale posizionale - o pochi tipi universali di segnali - e che le differenze nel corredo genetico o nella storia del precedente sviluppo delle cellule che devono reagire, determinino lo schema che si sviluppa realmente. Se è così, ne consegue che per lo sviluppo di disegni non circolari le cellule delle squame situate in differenti punti dell'ala devono seguire regole differenti per interpretare il segnale posizionale che proviene dal fuoco.

Possiamo visualizzare parecchi processi che producono disegni a simmetria bilaterale intorno a un fuoco. Attualmente prediligo un modello che richiede l'interazione di due gradienti; tale modello ha i vantaggi di essere visivamente semplice e, fatto ancora più importante, facilmente dimostrabile.

Per comprendere il modello è necessario formarsi un'immagine più concreta del segnale di informazione posizionale che emana dal fuoco. Provate a immaginarlo come una sostanza chimica, *P*, sintetizzata nel fuoco e che si diffonde allo

I disegni non circolari si sviluppano spesso intorno a un fuoco (una fonte d'informazione posizionale per le cellule che producono il pigmento). Dall'alto in basso sono rappresentate *Euphydryas editha*, *Hyalophora cecropia* e *Rothschildia arethusa*. Lo sviluppo dei disegni indica che il segnale che proviene dal fuoco non viene interpretato nello stesso modo da tutte le cellule circostanti.

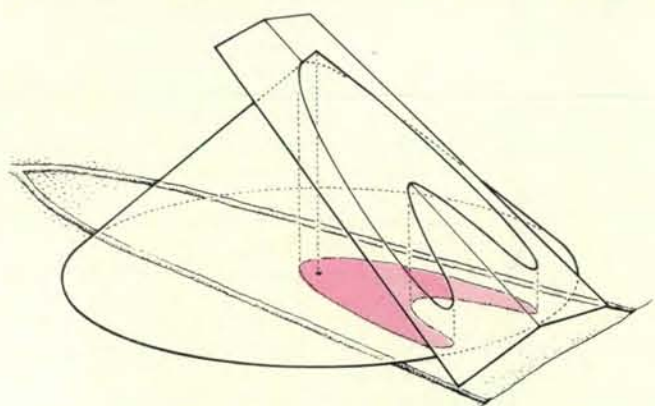
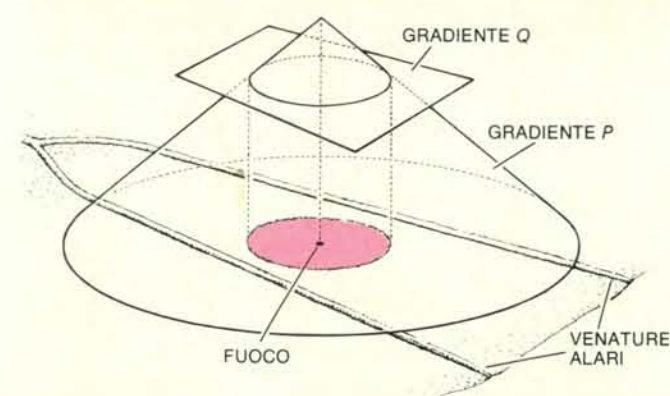
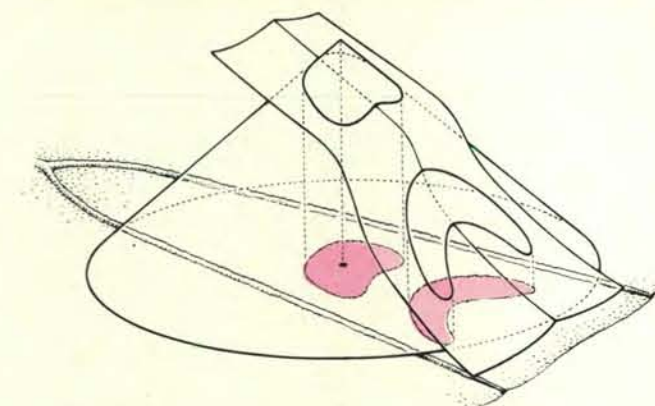
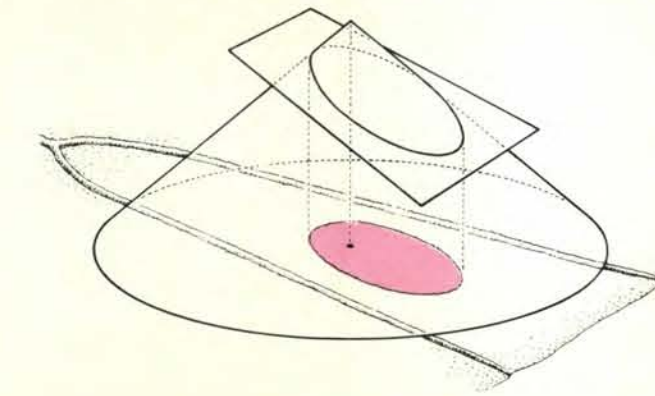
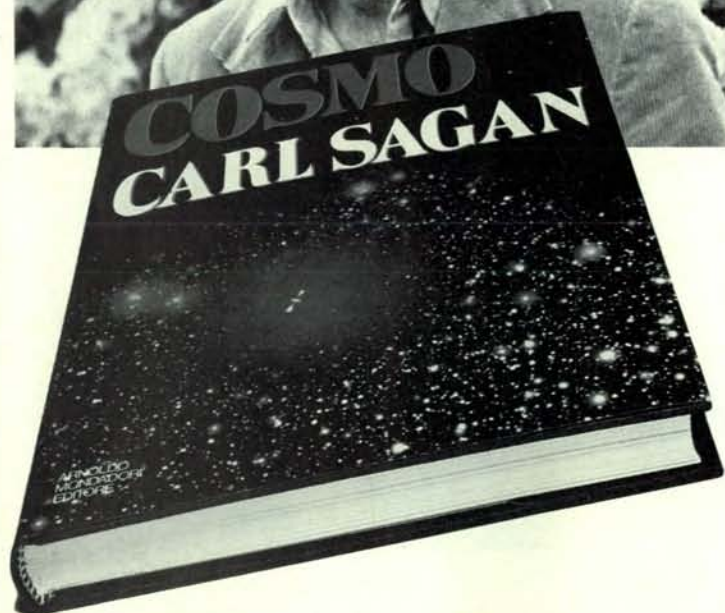
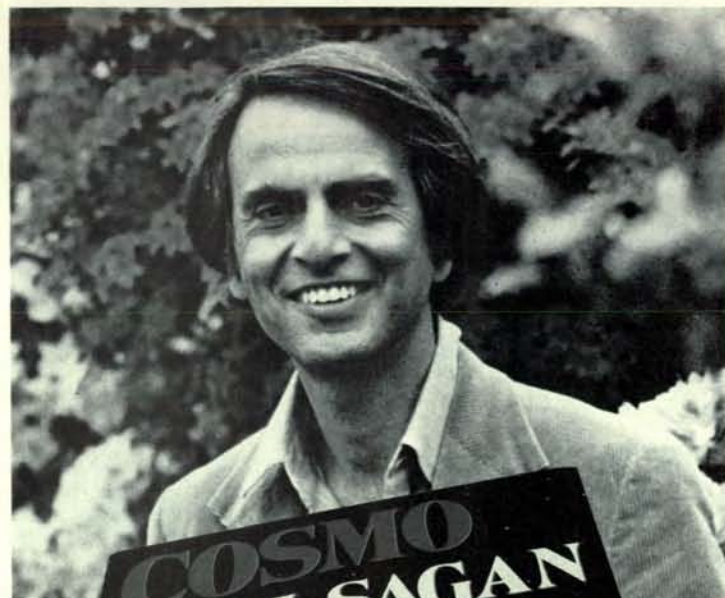
Lo schema di colorazione delle ali di farfalla è un mosaico di disegni indipendenti, uno per cellula alare (una cellula alare è una zona delimitata da venature). Il disegno completo di *Smyrna blomfildia*, una farfalla dell'America centrale, è in alto, quello delle singole cellule alari in basso. Anche i disegni sulle superfici superiore e inferiore dell'ala si sviluppano in modo indipendente.

stesso modo in tutte le direzioni. Il gradiente di concentrazione tridimensionale di P sarebbe un cono. La concentrazione più elevata si troverebbe nel fuoco e decrescerebbe gradualmente all'aumentare della distanza dal fuoco. La forma esatta del gradiente dipende da molti fattori, ma dal momento che esso non influisce sul fenomeno, considero il gradiente lineare.

Supponiamo che le cellule delle squame siano programmate per sintetizzare un pigmento nero dovunque P si trovi a una particolare concentrazione critica. Il risultato sarebbe la sintesi di un anello di pigmento nero. La situazione può essere rappresentata nel modello intersecando il gradiente a forma di cono di P con un piano orizzontale e proiettando la sezione conica così ottenuta sulla superficie dell'ala. L'altezza del piano corrisponde alla concentrazione critica a cui il pigmento nero viene codificato. Pigmenti differenti si possono rappresentare per mezzo di piani differenti e la larghezza di un anello di pigmento è proporzionale allo spessore del piano corrispondente. I piani sono dunque rappresentazioni grafiche delle regole secondo cui il gradiente di informazione posizionale viene interpretato.

Per quanto riguarda i disegni che si differenziano dalla forma circolare, dato lo stesso gradiente di informazione posizionale, le cellule delle squame poste su un lato del fuoco devono essere programmate per sintetizzare un dato pigmento a una concentrazione di P superiore a quella delle cellule delle squame situate sul lato opposto. Se un piano orizzontale (nella rappresentazione tridimensionale) significa che il gradiente di P viene interpretato allo stesso modo in qualunque punto dell'ala, una utile regola dice che l'altezza di una superficie di intersezione è proporzionale alla concentrazione di P a cui un certo pigmento viene sintetizzato. Un'ellisse sarebbe dunque generata da un piano obliquo.

A questo punto si è portati a pensare che un tale piano di intersezione sia un gradiente di una seconda sostanza chimica, Q, sintetizzata vicino alla base della cellula alare e che diffonde distalmente. È dunque possibile che tale pigmento venga sintetizzato solo quando esiste una particolare proporzione critica tra le concentrazioni delle sostanze chimiche. Con questo concetto non sarebbe necessario avere differenti sostanze chimiche per specificare la posizione di ogni pigmento nello schema; una diversa proporzione critica delle medesime due sostanze raggiungerebbe lo stesso effetto.

Sfortunatamente non si sa pressoché nulla sulle basi fisiologiche dello sviluppo dei disegni non circolari, tranne il fatto che in esso sono coinvolti i fuochi. Attualmente l'esistenza di un secondo gradiente è puramente ipotetica, ma tale ipotesi fornisce una direzione per future ricerche. Se un gradiente di questo tipo esiste, sarebbe possibile manipolarlo eliminandone chirurgicamente la fonte; dovrebbe essere anche possibile esplorare con fuochi trapiantati diverse regioni

Il meccanismo mediante il quale si generano gli ocelli marginali è un modello a due gradienti. Sotto ai disegni alari delle quattro specie si trova un diagramma in cui i due gradienti, P e Q, sono raffigurati come un piano e un cono intersecantisi. La forma tridimensionale di un gradiente lineare (una sostanza chimica che si diffonde allo stesso modo in tutte le direzioni all'interno di una cellula alare) sarebbe un cono, in cui l'origine della sostanza chimica, in questo caso P, si trova nell'apice. Q può essere lineare o non lineare. La sezione conica che nasce dall'intersezione dei piani non è il disegno sull'ala, che nasce dalla proiezione della sezione conica sull'ala. I disegni sono di complessità crescente e rappresentano da sinistra a destra: *Morpho peleides*, *Smyrna blomfildia*, *Dichorragia nesimachus* e *Speyeria aphrodite*.

del campo di questo secondo gradiente e vedere se i risultati confermano le previsioni sulla forma dei disegni che si genererebbero intorno a un fuoco trapiantato «a monte» o «a valle».

In molte specie di lepidotteri gli ocelli marginali hanno perso ogni rassomiglianza con cerchi o ellissi chiusi e possono assomigliare a punte di freccia, a galloni o a mezzelune, spesso con il lato convesso diretto dalla parte del fuoco e sempre con simmetria bilaterale rispetto alla linea mediana della cellula alare. Tali disegni si possono spiegare con il modello del doppio gradiente, se si suppone che almeno un gradiente non sia lineare e che il gradiente Q abbia una cresta o un punto di massimo che corre lungo la linea mediana e che si abbassa simmetricamente verso le venature alari. Ciò che questo modello in realtà propone è che il controllo della determinazione degli schemi di colorazione richieda almeno due fattori e che questi due fattori abbiano una distribuzione forse uguale a quella dei relativi valori di P e di Q.

Gli esperimenti pubblicati nel 1926 da Kühn offrono una convalida indiretta a un modello di questo tipo: egli ha esposto pupe della farfalla europea *Argynnis paphia*, della famiglia dei ninfalidi, a temperature vicine al congelamento durante il periodo in cui si determina lo schema di colorazione. Gli effetti fisiologici esatti di tali shock termici non sono noti, ma il loro risultato è stato un apparente arresto degli schemi di colorazione a vari stadi di sviluppo. La sequenza di schemi così ottenuta si può spiegare per mezzo del modello dei due gradienti se si considera che vi è un aumento progressivo del livello di Q accompagnato da una diffusione graduale del gradiente a forma di cono di P durante gli stadi precoci dello sviluppo degli schemi, quando i gradienti si stabiliscono per la prima volta.

Non sembrano necessari fattori addizionali per spiegare le migliaia di differenti disegni alari basati sul modello base dei ninfalidi. Pare che i singoli schemi specie-specifici nascano solo dalla manipolazione durante lo sviluppo di poche variabili: il numero e la localizzazione esatta dei fuochi, la forma relativa dei due gradienti, uno dei quali generato dal fuoco e l'altro dalla regione prossimale della cellula alare, la scelta dei pigmenti in cui si esprime il disegno e l'importanza dei segnali posizionali inerenti ai due gradienti che codificano i vari pigmenti. Di tutti questi fattori, l'unico attualmente noto è l'esistenza dei fuochi. I pochi di noi che lavorano in questo campo sono ancora ben lontani dalla comprensione della natura dei gradienti posizionali e dei meccanismi per mezzo dei quali essi influenzano la sintesi dei pigmenti specifici nelle squame dei lepidotteri. Comunque ora abbiamo almeno uno schema concettuale all'interno del quale possiamo iniziare a porci delle domande pratiche non solo sulla genesi dei singoli schemi di colorazione, ma anche sulla fisiologia comparata e sull'evoluzione di schemi biologici in generale.

Le fasi fluide della materia

Dopo un secolo di tentativi, è stata costruita una teoria attendibile della struttura dei liquidi e dei gas, basata su uno dei primi e più semplici modelli, nel quale le molecole sono sfere rigide e inerti

di J. A. Barker e Douglas Henderson

La differenza tra un liquido e un gas è ovvia nelle condizioni di temperatura e di pressione che si trovano comunemente sulla superficie terrestre. Un liquido può essere tenuto in un recipiente aperto riempiendolo fino al livello di una superficie libera. Un gas non forma alcuna superficie libera, ma tende invece a diffondere in tutto lo spazio disponibile; esso pertanto deve essere tenuto in un recipiente chiuso o deve essere imprigionato da un campo gravitazionale, come nel caso dell'atmosfera di un pianeta. Questa distinzione è stata uno degli aspetti salienti delle prime teorie che descrivevano le fasi della materia. Nel diciannovesimo secolo, per esempio, una teoria sosteneva che un liquido poteva venire «dissolto» in vapore senza perdere la propria identità, mentre un'altra teoria affermava che le due fasi sono costituite da diversi tipi di molecole: liquidoni e gasoni. Le teorie oggi più accreditate seguono un metodo alquanto diverso, evidenziando ciò che i liquidi e i gas hanno in comune: sono entrambi forme di materia prive di struttura permanente ed entrambi scorrono facilmente. Sono fluidi.

La fondamentale somiglianza tra i liquidi e i gas appare chiaramente evidente quando si fanno aumentare lievemente la temperatura e la pressione. Supponiamo di riscaldare un recipiente chiuso riempito parzialmente con un liquido. Il liquido espande o, in altri termini, diventa meno denso; una parte di esso evapora. Al contrario, il vapore sovrastante la superficie liquida diventa più denso quando a esso si aggiungono le molecole evaporate. La dilatazione del liquido e la maggior densità del vapore fanno aumentare la pressione nel recipiente. Riscaldandoli ulteriormente, la densità del liquido continua a diminuire e quella del vapore continua ad aumentare, finché le due densità non sono uguali. La combinazione dei valori di temperatura e pressione alla quale le densità diventano uguali è detta punto critico. Al di sopra del punto critico, liquido e gas diventano indistinguibili; esiste una sola fase fluida indistinta di densità uniforme.

Sul significato delle differenze strutturali e dinamiche tra liquidi e gas indagò per primo negli anni settanta del secolo scorso il fisico olandese J. D. van der Waals. A quell'epoca esisteva già una accreditata teoria dei gas perfetti, gas per i quali le relazioni tra temperatura, pressione e volume seguono una legge particolarmente semplice. Van der Waals propose di considerare un liquido come un gas denso che poteva essere descritto con una modifica dell'equazione di stato di un gas perfetto.

Le teorie più recenti della struttura fluida che qui descriveremo si possono ricollegare alla ricerca di van der Waals. Il modello concettuale fondamentale su cui si basano le teorie è un fluido composto da sfere rigide. Le proprietà del modello possono essere esplorate mediante simulazioni eseguite con un calcolatore. Il modello può anche servire come base per la descrizione di fluidi reali con metodi perturbativi, nei quali il modello fondamentale viene sempre più perfezionato aggiungendovi una successione di modificazioni sempre più raffinate.

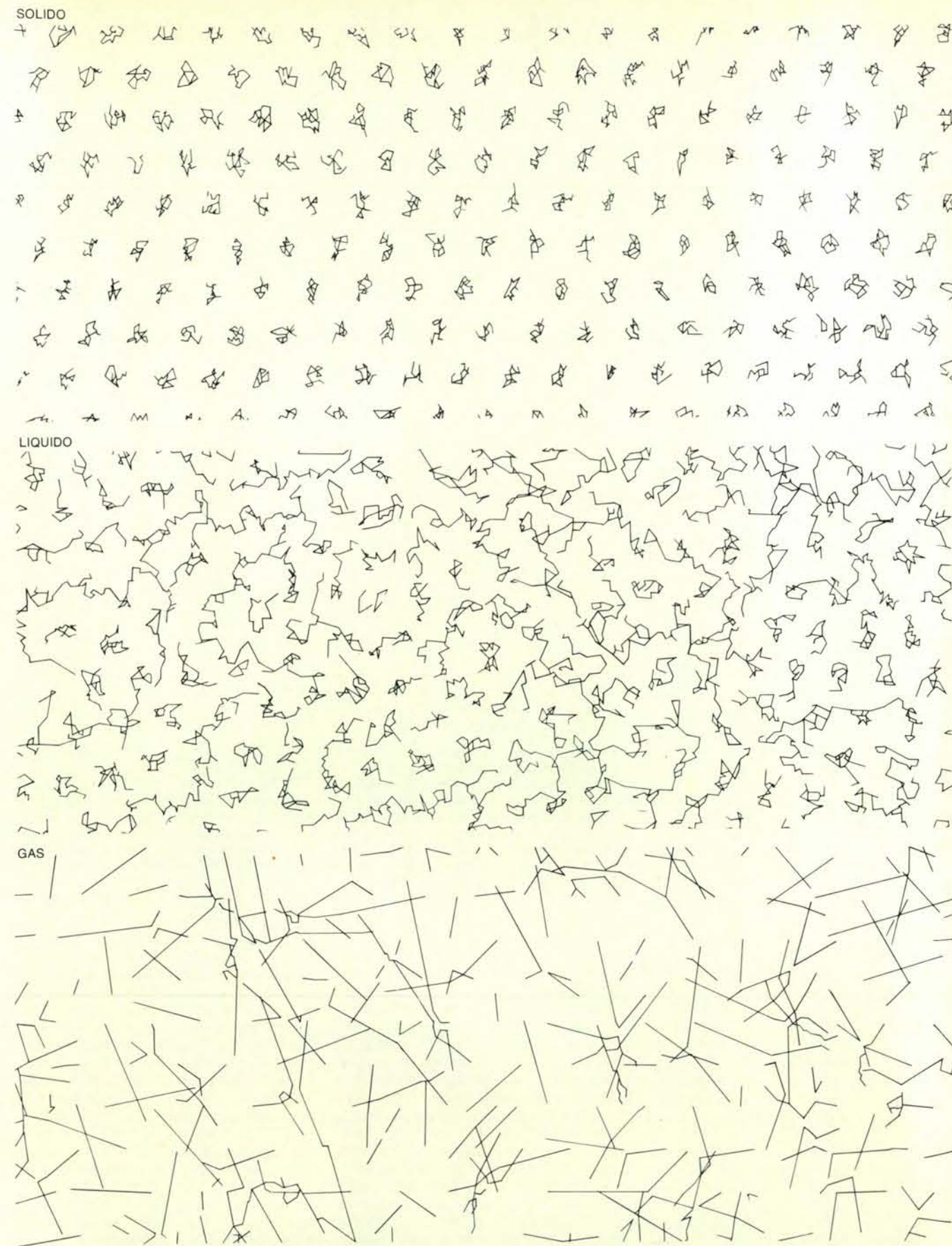
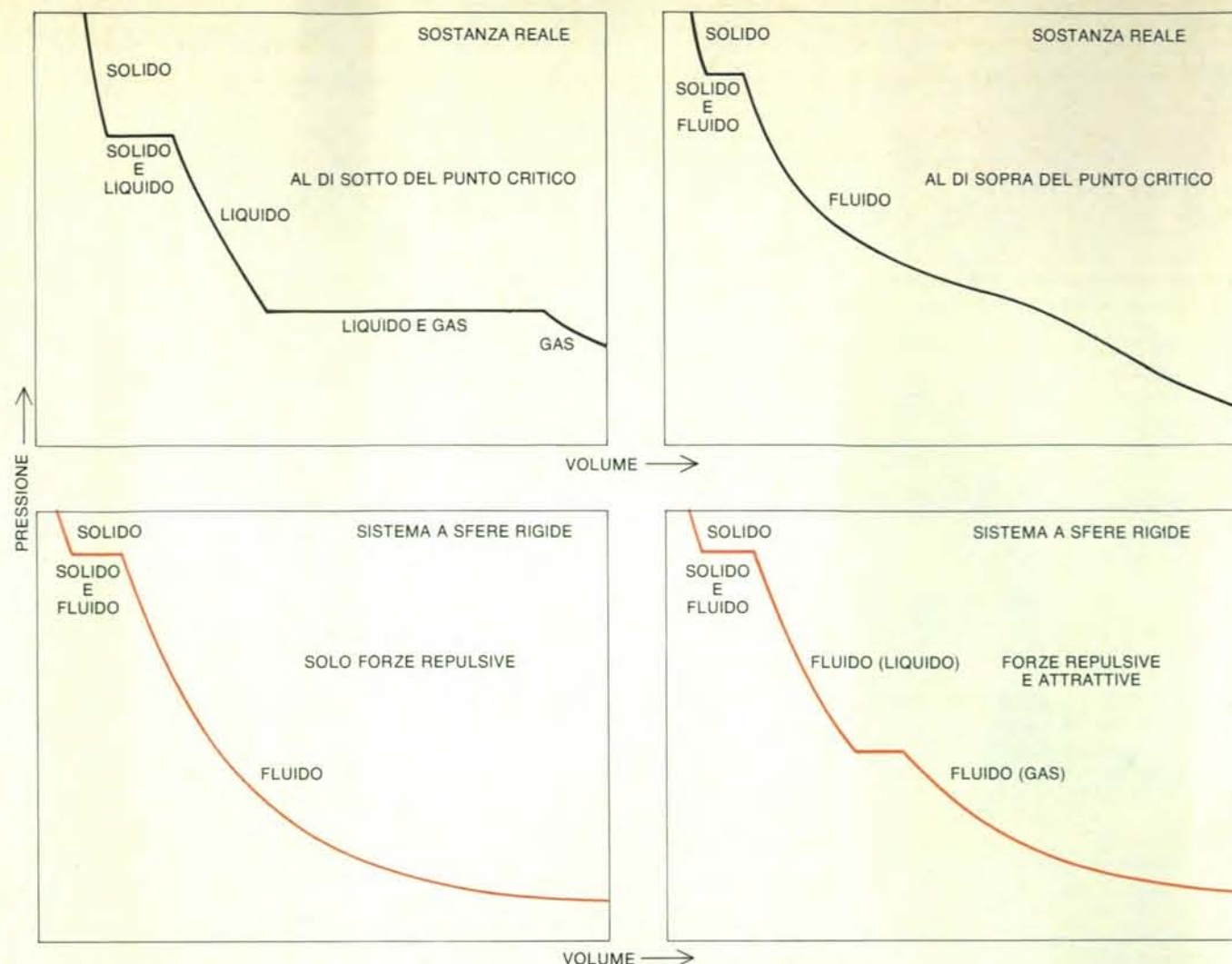
Virtualmente tutti gli elementi e tutti i composti chimici hanno una fase solida, una fase liquida e una di vapore. Una transizione da una fase a un'altra è accompagnata (o può essere causata) da una variazione di temperatura, pressione, densità o volume. La natura di tali transizioni può essere illustrata considerando ancora una sostanza in un recipiente chiuso, nel quale si possono controllare le condizioni che determinano la sua fase.

Supponiamo che la sostanza si trovi inizialmente allo stato solido e che la temperatura venga mantenuta costante mentre aumenta il volume del recipiente. In un primo tempo il volume del solido aumenta solo leggermente, e quindi la pressione diminuisce. Però, quando il solido comincia a fondere, la pressione diventa improvvisamente costante perché il volume totale occupato dal solido e dal liquido coesistenti va di pari passo con l'aumento di volume del recipiente. Quando tutta la sostanza è stata trasformata in liquido, la pressione ricomincia a diminuire, anche se non con la stessa rapidità che per il solido. Dopo un ulteriore aumento di volume della camera si raggiunge il punto di ebollizione, che è caratterizzato da un altro tratto a pressione costante, dato che la formazione del vapore compensa l'aumento di volume del recipiente. Infine, quando tutto il liquido è evaporato, la pressione riprende ancora una volta a scendere.

Se questo esperimento venisse eseguito a una temperatura superiore al punto critico della sostanza, non vi sarebbe alcuna transizione discontinua da liquido a vapore. Anzi, la pressione e la densità del fluido diminuirebbero in modo regolare. Si deve tuttavia far notare che la transizione solido-fluido continua fino alle più alte temperature alle quali le molecole mantengono la loro identità. Non esiste alcuna temperatura critica per la fusione.

L'obiettivo di una teoria che sia in grado di descrivere le fasi della materia è quello di spiegare osservazioni come queste in termini del moto dei singoli atomi o delle singole molecole che formano la sostanza. In altre parole, si deve stabilire un collegamento tra le proprietà termodinamiche di sostanze comuni e la meccanica delle singole molecole.

Nella maggior parte dei casi il moto delle molecole può essere descritto con precisione accettabile con i metodi della meccanica classica o newtoniana. Nella meccanica classica l'accelerazione di una molecola (cioè la variazione della sua velocità) è determinata interamente dalla sua massa e dalle forze su essa agenti. Se si conoscono in un dato istante le posizioni e le velocità di tutte le molecole in un campione di materia e se è nota la natura delle forze intermolecolari, allora nella meccanica classica si possono calcolare le posizioni e le velocità in tutti gli altri istanti, sia nel passato, sia nel futuro. Eseguendo un calcolo del genere, si potrebbero individuare in tutti i dettagli gli eventi che costituiscono un cambiamento di fase.

Un campo per cui la meccanica classica non può fornire una descrizione adeguata è la determinazione delle forze intermolecolari. Le sole forze significative sono

Le traiettorie delle molecole in un solido, in un liquido e in un gas sono state simulate con l'aiuto di un calcolatore. Il sistema bidimensionale di molecole ha le stesse fasi e le stesse transizioni di fase di una sostanza reale, ma le posizioni e i moti molecolari sono più facilmente visualizzati. Nel solido le molecole sono costrette a vibrare attorno a punti reticolari fissi, mentre le molecole delle fasi liquida e gassosa sono libere di spaziare. Le sole differenze sostanziali esistenti tra i due stati fluidi sono la densità e la frequenza di collisione. La simulazione è stata eseguita da Farid Abraham dell'International Business Machines Corporation Research Laboratory di San Jose in California.

La pressione e il volume di una sostanza reale e di un sistema modello composto da sfere rigide variano bruscamente al variare delle fasi. I diagrammi in alto mostrano la relazione tra pressione e volume per la sostanza reale a due temperature. Al di sotto della temperatura chiamata punto critico si osservano due transizioni di fase, ciascuna delle quali è rappresentata da un «piano», ossia da un segmento orizzontale nel diagramma pressione-volume. In queste regioni la sostanza si suddivide in due fasi distinte, e una leggera variazione di pressione o di temperatura può dare origine a una brusca transizione di fase. Al di sopra della temperatura critica la transizione solido-fluido persiste, ma liquido e gas sono indistinguibili; esiste soltanto un fluido supercritico il cui volume può essere modificato con continuità senza variazione di fase. I diagrammi in basso mostrano la relazione pressione-volume per il sistema a sfere rigide, in un caso, quando tra le sfere agiscono soltanto forze repulsive, e in un altro caso, quando si aggiungono forze attrattive. Nel sistema a sfere rigide l'aggiunta di forze attrattive crea un piano molto simile a quello che separa il liquido dal gas a temperature inferiori a quella critica. Pare che la formazione dello stato liquido richieda forze attrattive.

quelle elettromagnetiche; esse derivano da una distorsione che si sviluppa nella nube di elettroni che circonda una molecola, quando un'altra molecola si trova nelle vicinanze. Dato che un elettrone è molto più leggero di una molecola e si muove quindi più velocemente, la configurazione della nube di elettroni può essere descritta soltanto con i metodi della meccanica quantistica. I calcoli basati sulla meccanica quantistica sono ben più difficili di quelli basati sulla meccanica classica, ma si possono adottare due importanti semplificazioni nel definire le forze intermolecolari. In primo luogo, proprio perché l'elettrone si muove molto più velocemente dei nuclei atomici di una molecola, si può supporre che i nuclei siano in quiete. In secondo luogo, risulta che la forza tra due molecole è quasi indipendente dalla presenza di altre molecole. Pertanto dalla forza agente tra due sole molecole in quiete si può ricavare una buona approssimazione di tutte le forze agenti all'interno di un sistema costituito da molte molecole (anche molecole in moto rapido).

Anche l'obiettivo, notevolmente ridotto, di definire la natura della forza tra due molecole stazionarie si è dimostrato considerevole. In effetti, finora è stato raggiunto con precisione solo per i gas inerti, le cui molecole sono in realtà atomi singoli con una struttura elettronica particolarmente semplice e una simmetria sferica. L'importanza di quest'ultima proprietà sta nel fatto che essa indica come la forza tra le molecole sia la stessa in tutte le direzioni e dipenda soltanto dalla distanza tra esse. Una forza che dipende anche dall'orientazione delle molecole può dare origine a interessanti proprietà nelle fasi condensate della materia, come accade, per esempio, nell'acqua. Lo studio di tali forze anisotrope è al momento attuale un'area attiva di ricerca, ma noi concentreremo la nostra attenzione principalmente sulle forze isotrope, quali quelle dei gas inerti.

La forza tra due atomi di un gas inerte si può determinare in diversi modi. Per l'elio, nel quale il sistema di due atomi comprende soltanto quattro elettroni, le equazioni quantomeccaniche che definiscono la configurazione della nube elettronica si possono risolvere numericamente con un grande calcolatore. Un calcolo siffatto è stato eseguito dai nostri colleghi Bowen Liu e A. Douglas McLean dell'International Business Machines Corporation Research Laboratory di San Jose in California. Per i gas inerti più pesanti e per le molecole più complesse, i calcoli completi non sono di solito eseguibili, ma è possibile misurare sperimentalmente le forze intermolecolari. Alcune misurazioni sono del tutto indirette: la forza viene ricavata da certe proprietà dei gas che dipendono da collisioni bimolecolari, quali la viscosità e la conducibilità termica. Recentemente si sono resi disponibili metodi più diretti, nei quali si fanno scontrare fasci atomici o si analizza la radiazione assorbita ed emessa da molecole biatomiche debolmente legate come quelle dei gas inerti.

I risultati di tali calcoli o misurazioni vengono solitamente espressi mediante una funzione energia potenziale, dalla quale si può facilmente determinare la forza tra due molecole. La funzione precisa l'energia associata all'interazione delle molecole in relazione alla distanza tra esse. Le molecole tendono ad avvicinarsi a qualunque distanza renda minima la loro energia totale, e tale tendenza appare come una forza che spinge le molecole verso la distanza ideale. In un grafico della funzione energia potenziale la forza intermolecolare è espressa dall'opposto della pendenza della curva; nel punto di minima energia la pendenza è nulla e quindi non vi è alcuna forza. Alla distanza ideale l'energia dell'interazione è negativa, cioè l'energia del sistema composto dalle due molecole è inferiore a quella delle due molecole separate.

Grosso modo le funzioni energia potenziale per quasi tutte le molecole hanno la stessa forma fondamentale. Quando le molecole si trovano a distanza infinita non interagiscono affatto e quindi sia l'energia potenziale sia la forza tra le molecole vengono poste uguali a zero. Quando le molecole si avvicinano, le loro nubi elettroniche vengono distorte in modo tale che l'energia totale diminuisce leggermente; come conseguenza si ha una forza attrattiva che tende ad avvicinare ulteriormente le molecole. Questo processo continua finché le molecole distano una dall'altra all'incirca un diametro molecolare. Per ridurre ulteriormente la loro distanza, si deve spendere energia perché le nubi elettroniche cominciano a sovrapporsi; la forza diventa allora fortemente repulsiva, rispecchiando le interazioni repulsive degli elettroni. Il punto di equilibrio di minima energia e di forza nulla si raggiunge quando le interazioni attrattive e quelle repulsive sono in perfetto equilibrio.

La semplicità della funzione energia potenziale per una coppia di molecole pone immediatamente un problema: come è possibile che un'interazione così semplice tra molecole dia origine alle tre fasi della materia con le loro proprietà caratteristiche? La funzione ha soltanto un punto di minima energia. In un solido le molecole sono disposte in modo tale che la loro distanza sia il più vicino possibile a questa distanza di minima energia. Allora che cosa giustifica l'esistenza dei liquidi e dei gas? Dal momento che la loro energia deve essere evidentemente maggiore di quella minima, perché essi non diminuiscono la loro energia condensandosi in solidi? Ci si potrebbe aspettare che la funzione energia potenziale abbia un secondo e un terzo minimo per sistemare le fasi fluide, ma questi minimi ulteriori non esistono.

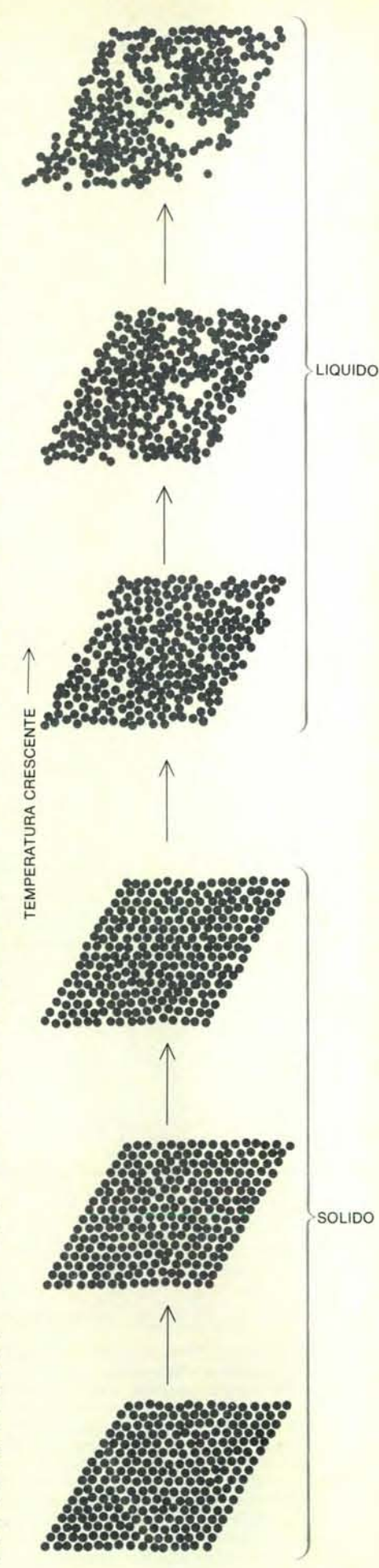
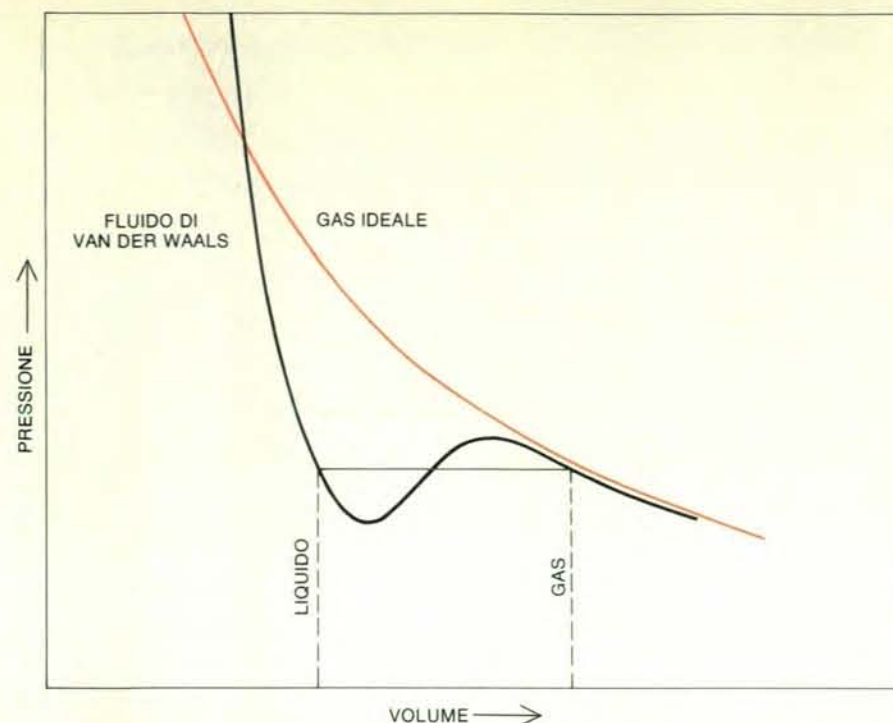
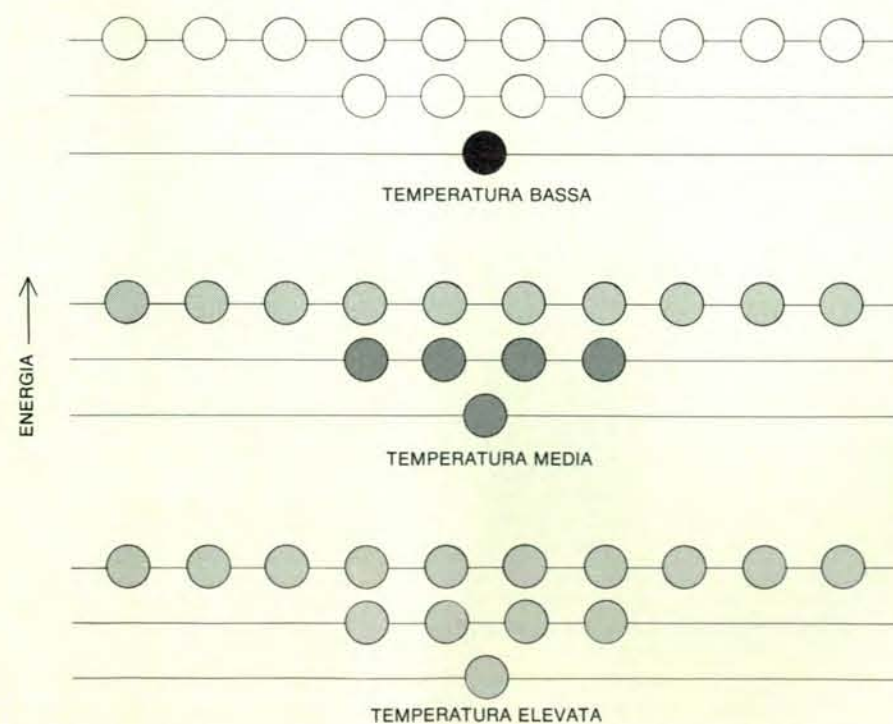
La soluzione di questo paradosso sta nel fatto che il sistema a molte molecole, in un campione di materia macroscopico, non è un semplice sistema meccanico, e l'energia non è la sola grandezza ad avere effetto sul suo comportamento. Il sistema di molecole è retto dai principi della termodinamica e della meccanica statistica, dove è assegnato un ruolo importante anche alla temperatura e alla misura del disordine chiamata entropia.

Un semplice sistema meccanico, quale una sfera che rotola giù per un pendio, raggiunge normalmente le condizioni di quiete nel suo stato di minima energia, cioè al fondovalle. Allo stesso modo, due molecole isolate tendono sempre ad assumere la loro distanza di minima energia. Un sistema termodinamico, invece, si avvicina non allo stato di minima energia, ma allo stato di minima energia libera. L'energia libera *(A)* è data dalla differenze tra l'energia *(E)* e il prodotto della temperatura *(T)* per l'entropia *(S)*; quindi un sistema termodinamico è retto dall'equazione $A = E - TS$.

Da questa relazione si possono dedurre le caratteristiche salienti di un sistema termodinamico. A bassa temperatura l'energia delle molecole dà il maggior contributo all'energia libera e pertanto lo stato favorito è la fase solida, caratterizzata dalla minima energia. A temperatura più elevata l'influenza predominante è quella dell'entropia del sistema. Il risultato è che le fasi fluide sono stabili a temperature elevate anche se esse costituiscono configurazioni a energie più elevate rispetto alla fase solida.

I concetti di entropia e di energia libera si possono comprendere considerando la relazione tra le descrizioni macroscopica e microscopica di un sistema termodinamico. Lo stato macroscopico di un sistema è individuato da grandezze quali la temperatura e la densità, che sono proprietà collettive di tutte le molecole. Lo stato microscopico è definito dalla configurazione particolareggiata delle molecole. È un fatto essenziale che a un solo stato macroscopico possano corrispondere molti stati microscopici. In altri termini, molte configurazioni delle molecole possono dare origine agli stessi valori di temperatura, densità e altre grandezze.

Nella meccanica statistica l'entropia è interpretata come una misura della probabilità intrinseca di uno stato macrosco-

La configurazione delle molecole di una sostanza bidimensionale è visibile qui in una serie di «istantanee» scattate durante una simulazione al calcolatore. Nell'immagine in basso la sostanza è un solido, che si espande e si «distorce» leggermente al crescere della temperatura. Quando si raggiunge la temperatura di fusione si verifica una transizione discontinua dallo stato solido allo stato fluido. Nel fluido le molecole si muovono liberamente, dando luogo a vuoti momentanei, man mano che la sostanza subisce una ulteriore espansione. La simulazione è stata effettuata da Abraham.

La teoria dei liquidi venne formulata da J. D. van der Waals modificando l'equazione di stato che definisce la relazione tra pressione, volume e temperatura per un gas perfetto. La buca nel potenziale di van der Waals è dovuta all'introduzione di forze attrattive tra le molecole e alla modifica del volume disponibile provocato dall'esclusione del volume proprio delle molecole. Le forze di attrazione intermolecolari hanno la funzione di ridurre la pressione totale del fluido.

La configurazione microscopica di un sistema di molecole determina lo stato macroscopico del sistema. Viene qui indicato in modo approssimato il numero relativo delle configurazioni possibili mediante il numero di cerchietti su ciascuno dei tre livelli energetici. La probabilità che il sistema abbia una particolare configurazione è indicata dall'intensità di annerimento del cerchio. La probabilità è determinata sia dall'energia della configurazione, sia dalla temperatura del sistema. A bassa temperatura la configurazione di minima energia è la sola dotata di un'apprezzabile probabilità; tale configurazione corrisponde allo stato solido. A temperatura intermedia sono favorite le configurazioni a bassa energia, ma, dato che esistono molte più configurazioni ad alta energia, ciascun livello energetico può avere all'incirca la stessa probabilità totale; tale situazione prevale durante la transizione da fase solida a liquida. Ad alta temperatura le probabilità di tutte le configurazioni sono pressoché identiche. Si è quasi sicuri di osservare una configurazione ad alta energia perché tali configurazioni sono più numerose. Il solido ha subito la fusione.

pico. La probabilità è determinata dal numero di stati microscopici che corrispondono allo stato macroscopico e dalla probabilità assegnata a ciascuno degli stati microscopici. Tutti gli stati microscopici con la stessa energia hanno la stessa probabilità, ma la probabilità di uno stato diminuisce rapidamente (secondo una curva esponenziale) all'aumentare dell'energia di uno stato o al diminuire della temperatura. L'espressione matematica che definisce la dipendenza della probabilità dall'energia e dalla temperatura è detta «fattore di Boltzmann», dal nome del fisico austriaco Ludwig Boltzmann.

Il fattore di Boltzmann implica che i singoli stati microscopici che hanno la minima energia siano sempre quelli più probabili, ma non ne deriva che lo stato macroscopico osservato abbia necessariamente una bassa energia. Alcuni stati macroscopici di energia più alta possono essere prodotti da un numero molto maggiore di stati microscopici di alta energia. Anche se ciascuno degli stati microscopici di alta energia ha una bassa probabilità, gli stati possono essere talmente numerosi che il sistema ha una notevole probabilità di occupare l'uno o l'altro di essi. Con questa interpretazione i fluidi esistono perché vi è un maggior numero di modi di organizzare le molecole in un liquido che in un solido e ve ne è un numero ancora maggiore in un gas.

La teoria della meccanica statistica è stata presentata in forma completa dal fisico americano Josiah Willard Gibbs nel suo libro *Elementary Principles in Statistical Mechanics*, pubblicato nel 1902. Gibbs dimostrò come un sistema di particelle a una data temperatura si distribuisca tra i vari livelli energetici disponibili. I suoi eleganti risultati colmano il vuoto tra due grandi visioni del mondo: quella meccanica e quella termodinamica.

Con una conoscenza delle forze intermolecolari e della meccanica statistica erano a disposizione gli strumenti concettuali fondamentali necessari per studiare le fasi della materia; mancavano tuttavia ancora alcuni strumenti pratici. Prima della metà del ventesimo secolo vi furono soddisfacenti teorie dei solidi, che sono altamente ordinati, e dei gas, che sono solo moderatamente densi e altamente disordinati. Non esisteva però alcuna teoria pienamente soddisfacente per i liquidi e i gas compresi.

Il principale ostacolo allo sviluppo di una tale teoria era il grande numero di molecole che si doveva inserire in qualsiasi descrizione di un liquido o di un gas denso. Tali problemi a molti corpi non si possono risolvere analiticamente; in realtà, neppure il problema dei tre corpi ha una soluzione esatta tranne che in particolari circostanze. Per descrivere i fluidi densi si dovettero sviluppare metodi di approssimazione. Il metodo più diretto è la simulazione del fluido con un calcolatore.

Il primo passo nella costruzione di un modello matematico di un fluido è la scelta di un'appropriata funzione energia potenziale per definire la forza tra ciascuna coppia di molecole. Si può utilizzare una funzione ricavata dai calcoli della meccanica quantistica o dalle misurazioni sperimentali, ma è spesso vantaggioso disporre di una funzione con una struttura matematica più semplice. Per esempio, verso il 1930 John E. Lennard-Jones introdusse una semplice funzione che approssimava molto da vicino la forma della funzione che determina le interazioni tra gli atomi di gas inerti. Sono state eseguite molte ricerche esplorative con una funzione ancor più semplice, cioè quella del modello a sfere rigide. Nel fluido a sfere rigide non esiste alcun potenziale attrattivo e le sfere sono perfettamente rigide, in modo tale da non consentire alcuna interpenetrazione. Quindi il potenziale repulsivo è infinitamente ripido.

Un modello comprende un insieme di molecole, alle quali si devono assegnare le posizioni iniziali. In ogni fase dell'evoluzione del modello si determinano nuove posizioni che rappresentano un nuovo stato del sistema. In alcuni casi vengono determinate in ogni stadio anche le velocità di tutte le molecole. Dalle medie di tutte le configurazioni osservate dopo molte fasi, è possibile determinare le proprietà macroscopiche del fluido modello e confrontarle con i risultati sperimentali dei fluidi reali.

Un metodo per generare stati successivi del sistema modello è detto metodo di Monte Carlo perché il suo meccanismo essenziale è un generatore di numeri casuali la cui funzione è concettualmente equivalente a quella della ruota di una roulette. L'algoritmo su cui si basa il metodo è stato inventato nel 1953 da Nicholas C. Metropolis, Arianna W. e Marshall N. Rosenbluth e Augusta H. ed Edward Teller; è stato applicato nella sua formulazione primitiva da William W. Wood. L'algoritmo genera una successione di configurazioni di un sistema a molte particelle che imita alcune proprietà statistiche del moto molecolare. In particolare la probabilità che una data configurazione appaia nella successione è proporzionale al fattore di Boltzmann di tale configurazione, proprio come accade in un sistema termodinamico reale.

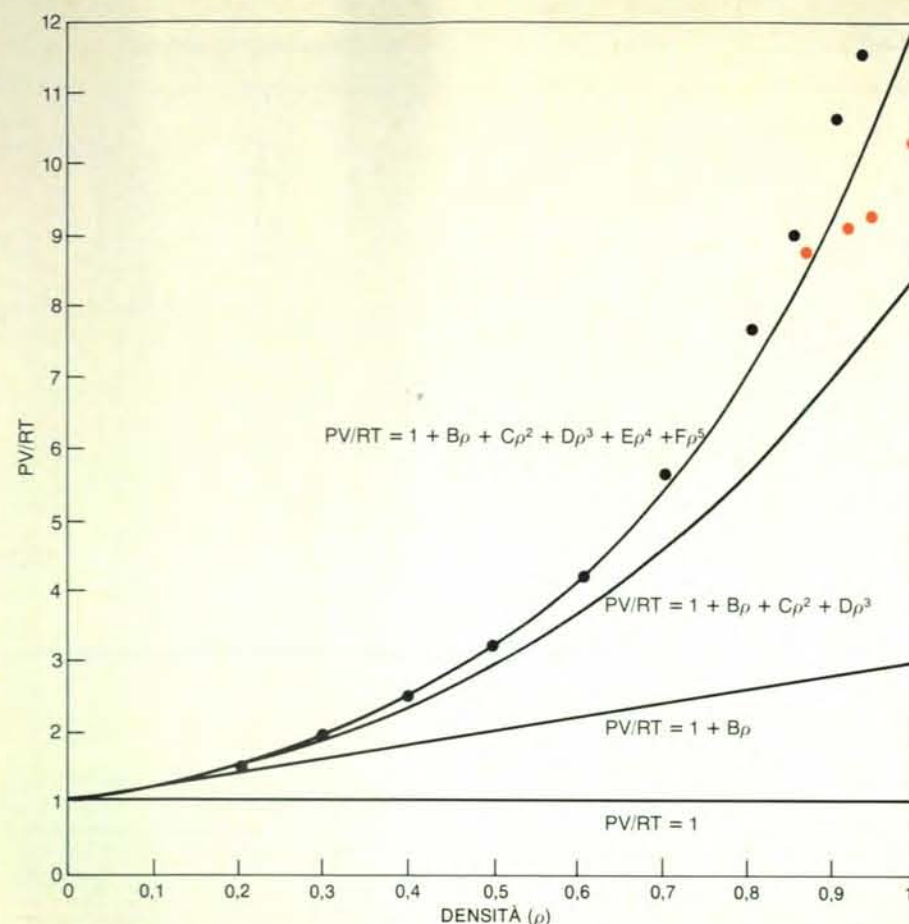
Il diametro efficace delle sfere rigide diminuisce all'aumentare dell'energia o della temperatura. A bassa energia le collisioni molecolari sono mediamente meno intense e si ha una trascurabile sovrapposizione delle nubi elettroniche. A energia più alta le collisioni molecolari sono più violente e le nubi elettroniche si sovrappongono di più. A questo punto i centri delle molecole si avvicinano maggiormente e il diametro efficace delle molecole è inferiore. Tale effetto può essere incorporato in un fluido a sfere rigide rendendo il diametro una funzione della temperatura.

Nel metodo di Monte Carlo il programma del calcolatore sposta una molecola alla volta in accordo con le regole della meccanica statistica. Tuttavia, a causa dello stretto addensamento delle molecole, uno spostamento può portare una molecola talmente vicina a una delle sue molecole adiacenti da farle sovrapporre. A quel punto le forze repulsive fanno diventare grande e positiva l'energia potenziale. Dal momento che il fattore di Boltzmann diminuisce esponenzialmente al crescere dell'energia, la sovrapposizione fa sì che il fattore di Boltzmann diventi in realtà nullo per la maggior parte delle configurazioni. In pratica tale difficoltà viene superata calcolando, per ogni configurazione di una serie, la variazione di energia provocata da ogni spostamento. Se una molecola si avvicina troppo a quelle adiacenti, il programma la riporta in realtà nella sua posizione originale. Con questa modifica la frequenza di ogni configurazione è proporzionale al proprio fattore di Boltzmann.

Un altro metodo di simulazione al calcolatore è stato sviluppato da B. J. Alder e T. E. Wainwright del Lawrence Berkeley Laboratory, da Aneesur Rahman dell'Argonne National Laboratory e da Loup Verlet del Laboratorio di fisica teorica e delle alte energie di Orsay in Francia. È chiamato metodo della dinamica molecolare e si basa su una soluzione diretta delle equazioni classiche del moto per un sistema di molecole. In ogni stadio dell'evoluzione del modello si ricorre alla funzione energia potenziale per calcolare la forza esercitata su ogni molecola del campione dalle molecole circostanti. Si ricava l'accelerazione dalla forza risultante e dall'accelerazione si calcolano la nuova posizione e la nuova velocità.

Diversamente dal metodo di Monte Carlo, il metodo della dinamica molecolare fornisce informazioni sull'evoluzione del sistema al passare del tempo. Esso perciò può essere impiegato per calcolare proprietà del sistema fluido dipendenti dal tempo, come la viscosità. Si possono ricavare anche proprietà come l'entropia mediando le configurazioni osservate in un certo intervallo di tempo.

Una descrizione di un fluido che si può ricavare da una simulazione al calcolatore è la funzione di distribuzione radiale, che fornisce la densità media delle molecole in funzione della distanza radiale da una molecola di riferimento. In un fluido modello il calcolo della funzione di distribuzione radiale è un procedimento immediato: basta tabulare le distanze tra una data molecola e tutte le altre molecole. La funzione di distribuzione calcolata per il modello si può poi confrontare con la funzione di un solido o di un fluido reale. La funzione di distribuzione radiale per una sostanza reale si determina sperimentalmente misurando gli angoli ai quali i raggi X o i neutroni vengono diffusi dagli atomi o dalle molecole.

La funzione di distribuzione radiale contraddistingue nettamente le fasi della materia. A piccole distanze la funzione è nulla in tutti i casi, poiché due molecole non possono occupare la stessa posizione. Al di là di questa zona vietata la funzione per un gas a bassa densità e ad alta temperatura è quasi uniforme su tutte le distanze, perché la distribuzione delle molecole è casuale. In un solido cristallino la fun-

Una serie di perturbazioni porta a una successione di descrizioni sempre più accurate di un fluido a sfere rigide. Il sistema di riferimento con cui ha inizio la serie è un gas perfetto. L'equazione di stato per il gas perfetto si può scrivere nella forma $PV/RT = 1$, dove P è la pressione, V il volume, R una costante e T la temperatura; pertanto per il gas perfetto il valore di PV/RT resta costantemente pari a 1 qualunque sia la densità. In un gas reale ad alta temperatura il valore di PV/RT tende ad aumentare con la densità. Tale aumento viene giustificato dalla teoria aggiungendo all'equazione di stato una serie di termini nei quali la densità viene elevata successivamente a potenze crescenti. In tal modo l'equazione assume la forma $PV/RT = 1 + B\varrho + C\varrho^2 + D\varrho^3$ e così via. La somma di sei termini del genere costituisce una buona approssimazione ai risultati dei calcoli eseguiti con il calcolatore sul modello a sfere rigide dagli autori *(punti neri)* e da B. J. Alder e T. E. Wainwright del Lawrence Berkeley Laboratory *(punti in colore)*. Tuttavia, per un modello dello stato liquido un gas perfetto non costituisce un soddisfacente sistema di riferimento. Alle densità dei liquidi e a bassa temperatura i termini successivi diventano maggiori, non minori.

zione di distribuzione presenta una serie di picchi netti.

Le funzioni di distribuzione radiali più interessanti sono quelle dei liquidi e dei gas densi. Un primo picco a una distanza di circa un diametro molecolare rappresenta le molecole più prossime. Dato che esiste un'alta probabilità che le molecole circostanti siano in contatto, le più prossime cadono all'interno di un ristretto intervallo di distanze e il picco è stretto. Dopo il primo massimo vi è un minimo della funzione, seguito da un massimo più piccolo a circa due diametri molecolari, che corrispondono alla formazione di un secondo strato di molecole. A distanza maggiore i picchi diventano più larghi e al di là di circa quattro diametri la modulazione della funzione di distribuzione sparisce del tutto; ivi la densità delle molecole si avvicina alla densità uniforme del liquido nel suo insieme.

La funzione di distribuzione radiale dell'argo è stata determinata con precisione con la diffusione di neutroni. I risultati delle simulazioni al calcolatore dell'argo liquido sono in eccellente accordo con i risultati sperimentali, sia per la funzione di distribuzione radiale, sia per le proprietà termodinamiche quali le relazioni tra pressione, volume e temperatura.

Sebbene il metodo di Monte Carlo e il metodo della dinamica molecolare impieghino algoritmi differenti, i dettagli operativi dei due modelli sono simili. Se si devono ultimare i calcoli in un intervallo di tempo ragionevole, il numero di molecole deve essere limitato. Nella maggior parte delle applicazioni il numero oscilla da circa 100 a circa 1000. Solitamente si suppone che le molecole riempiano una scatola le cui dimensioni sono determinate dalla densità del fluido. Con un numero alquanto piccolo di molecole nella scatola gli effetti superficiali assumono un'importanza molto forte ed evidente.

Il programma del calcolatore corregge parzialmente gli effetti superficiali attraverso l'uso di condizioni periodiche al contorno. La scatola è circondata da tutti i lati da molte copie periodiche della stessa, come le celle unitarie di un cristallo. Una molecola nella scatola può interagire con molecole «fantasma» nelle scatole copia. Inoltre, tutte le volte che una molecola esce dalla scatola, il suo fantasma entra automaticamente dalla faccia opposta. In questo modo viene mantenuto costante il numero di molecole.

È risultato che un sistema con solo 100 molecole può fornire stime notevolmente precise delle proprietà delle fasi condensate di sostanze semplici come l'argo. La sola regione del diagramma di fase in cui i modelli sono carenti è quella in prossimità del punto critico. Nei pressi del punto critico si sviluppano, su molte scale diverse, fluttuazioni di densità e di altre proprietà. Anche le interazioni a grande distanza, tra le quali l'interazione elettrostatica tra ioni e l'interazione tra molecole polari come l'acqua, presentano particolari difficoltà.

Una simulazione fatta con un calcolatore digitale rappresenta un tentativo di calcolare le proprietà di un fluido dai principi primi. La seconda importante linea di sviluppo della teoria dei fluidi segue una impostazione del tutto differente. In primo luogo si cerca un modello matematicamente trattabile che fornisca una descrizione del sistema ragionevolmente buona; successivamente si aggiungono a esso ulteriori perfezionamenti.

Metodi di questo genere, che si chiamano metodi di perturbazione o perturbativi, vennero applicati per la prima volta al calcolo delle orbite planetarie. In tale contesto l'approssimazione iniziale deve considerare isolati il Sole e i pianeti. Quindi si aggiungono all'approssimazione, una alla volta, le perturbazioni dell'orbita causate dalla presenza di altri pianeti. Le correzioni assumono solitamente la forma di una serie di potenze in cui una piccola grandezza x viene elevata a esponenti sempre maggiori. La valutazione della serie diventa sempre più difficile al crescere del numero dei termini ma, se x è piccolo, l'importanza dei termini diminuisce rapidamente quando x viene elevato a potenze più alte. Sono quindi necessari soltanto pochi termini per ottenere la precisione desiderata. La parte più difficile di questo metodo è la ricerca di una x che sia abbastanza piccola da fornire risultati precisi con un numero ragionevole di termini.

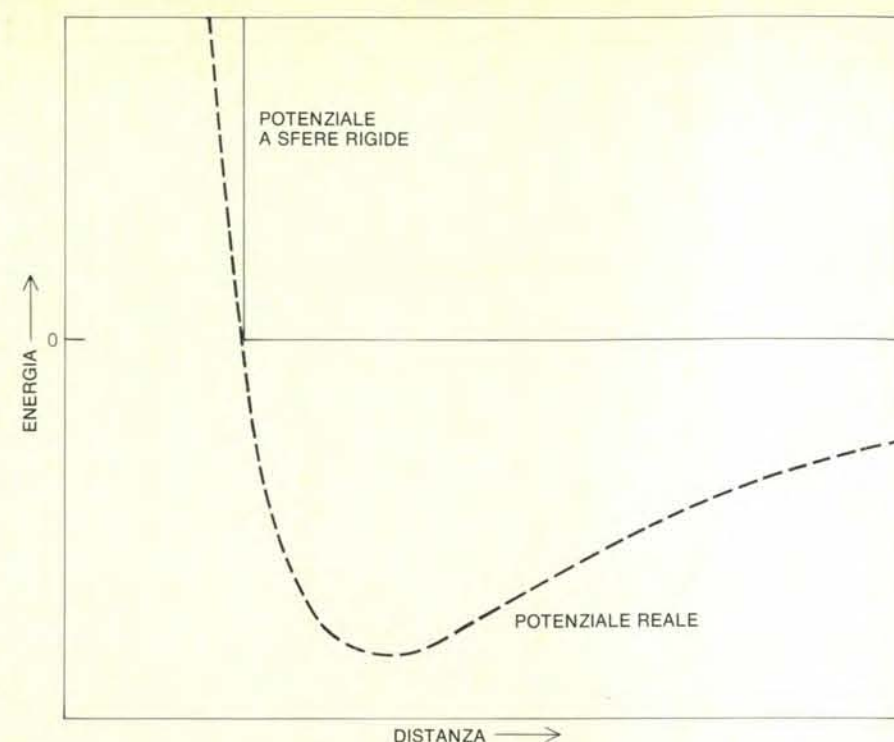
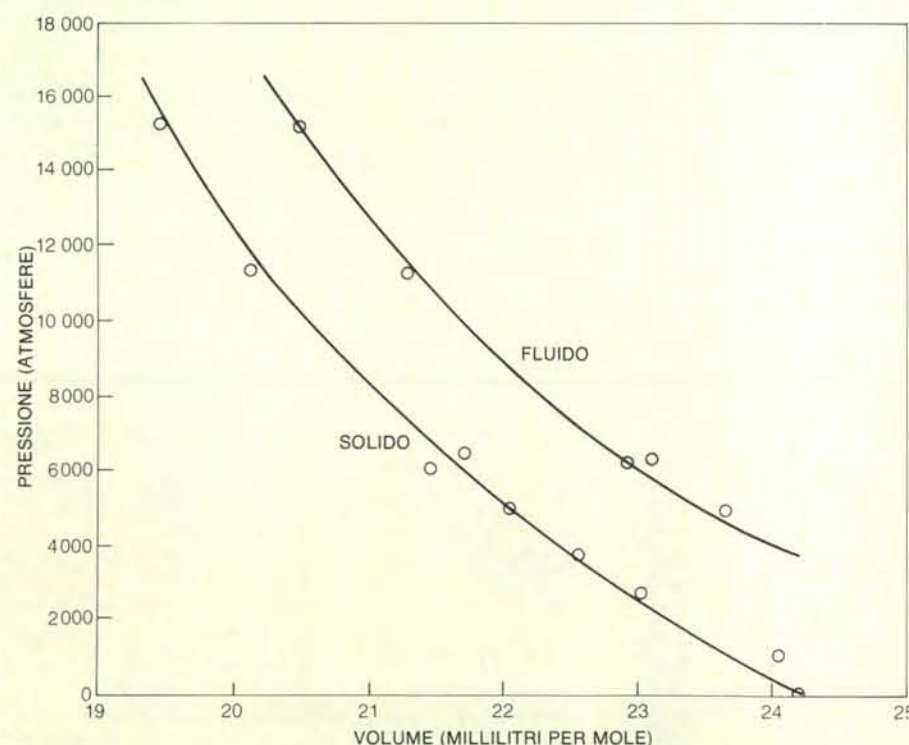
Nell'applicare i metodi perturbativi alla teoria dei fluidi vi è un ovvio candidato al ruolo di sistema di riferimento: si tratta del gas perfetto, costituito da molecole che non subiscono alcuna interazione. In tale gas la pressione è direttamente proporzionale alla temperatura e inversamente proporzionale al volume; la costante di proporzionalità è la costante dei gas, R. La relazione è espressa matematicamente dall'equazione $P = RT/V$, conosciuta come legge dei gas perfetti.

I metodi perturbativi suggeriscono un modo di modificare la legge dei gas perfetti per descrivere le relazioni tra pressione, temperatura e volume in un gas reale. Alla legge viene aggiunta una serie di termini in cui la densità del gas viene elevata a potenze successivamente crescenti. Questa teoria perturbativa dei gas imperfetti venne elaborata in dettaglio da Joseph E. Mayer. Essa fornisce una spiegazione soddisfacente fin quando la densità non diventa troppo grande. Ad alta densità i termini successivi della serie diventano maggiori anziché minori ed è quindi impossibile un calcolo preciso. Per questo motivo il gas perfetto non è un sistema di riferimento soddisfacente per una teoria dei liquidi, soprattutto perché le disposizioni molecolari importanti nel gas perfetto sono molto diverse da quelle di un liquido.

Un sistema di riferimento più promettente per i liquidi è il modello a sfere rigide. In questo modello le interazioni repulsive che diventano importanti ad alta densità sono giustificate in modo approssimativo e le disposizioni molecolari più comuni sono simili a quelle dei liquidi.

Il primo tentativo di descrivere i liquidi con un modello a sfere rigide fu eseguito da van der Waals. Egli propose di modificare la legge dei gas perfetti in due modi. In primo luogo, egli suggerì di sostituire il volume V con $V - b$, dove b (il covolume di van der Waals) rappresenta il volume dal quale una molecola è esclusa dalle interazioni repulsive con le altre molecole. Con l'ipotesi che le molecole siano sfere rigide si può calcolare un valore di b. La seconda modifica proposta fu quella di sottrarre dalla pressione una «pressione interna» definita dall'espressione a/V^2, dove a è una costante. È attraverso la seconda modifica che viene introdotto l'effetto delle forze attrattive. In tal modo l'equazione della pressione di un liquido diventa $P = RT/(V - b) - a/V^2$.

Van der Waals giustificò la sottrazione della pressione interna richiamando le sue idee intuitive sull'«associazione» di molecole. Quando una molecola si avvicina alla superficie di un contenitore, secondo van der Waals, essa ha più molecole dietro di sé che non davanti. La conseguenza è che non tutte le forze attrattive tra le altre molecole si annullano a vicenda. L'effetto è quello di creare una debole forza che riporta indietro la molecola nel volume del liquido e quindi diminuisce la forza esercitata dalle molecole sulle pareti del recipiente. La diminuzione della forza sulle pareti equivale a una diminuzione della pressione. L'effetto diventa apprezzabile soltanto nei liquidi e nei gas compressi, dove le molecole sono sufficientemente schiacciate una contro l'altra da far diventare significativa la forza attrattiva tra loro. La pressione interna è perciò una funzione della densità.

Sulla base di queste formulazioni van der Waals fu in grado di dimostrare che gli stati di liquido e di gas si scambiano con continuità l'uno con l'altro. Le sue idee generali sulla continuità degli stati e sull'addensamento delle molecole erano ben chiare; perfino le sue approssimazioni

La funzione energia potenziale determina la forza agente tra due molecole, che si manifesta come una tendenza delle molecole ad adattare le loro distanze a qualsiasi distanza che renda minima l'energia potenziale. Il potenziale per un sistema di sfere rigide è rappresentato da una linea verticale a una distanza corrispondente al diametro delle sfere e da una linea orizzontale a tutte le distanze superiori. La linea verticale indica che l'energia diventa infinita quando due molecole vengono schiacciate una contro l'altra a una distanza inferiore al loro diametro; in altri termini, una forza repulsiva infinita impedisce alle molecole di sovrapporsi. La linea orizzontale indica che non esiste alcuna forza tra le molecole quando sono a una distanza superiore al loro stesso diametro. Una funzione energia potenziale reale differisce dal potenziale di sfere rigide in due modi: la parte repulsiva del potenziale non è del tutto verticale e può aversi una certa sovrapposizione delle molecole; esiste, inoltre, una regione a energia potenziale negativa in cui la forza è attrattiva.

Le fasi solide e fluide di una sostanza possono entrambe essere descritte con esattezza da un modello nel quale una sola funzione energia potenziale definisce le interazioni delle molecole. Nel caso riportato in figura la relazione tra pressione e volume è rappresentata graficamente per l'argo; i cerchietti rappresentano i dati sperimentali, mentre le curve continue sono i risultati ottenuti per mezzo di una simulazione al calcolatore basata sulla funzione potenziale.

specifiche, quale l'espressione a/V^2 per la pressione interna, sono quasi esatte. Ciononostante van der Waals non presentò una teoria soddisfacente del fluido a sfere rigide. Tale teoria divenne abbordabile solo con le simulazioni al calcolatore di fluidi a sfere rigide, circa 20 anni fa.

Il punto debole della teoria dei fluidi di van der Waals era la grandezza *b*, che rappresentava il covolume. Se si suppone che *b* sia una costante, l'equazione di van der Waals fornisce solo una modesta approssimazione della pressione a sfere rigide. Ad alta densità l'approssimazione fa una stima grossolanamente superiore della pressione, perché i volumi esclusi dalle differenti molecole si sovrappongono in modo tale che il covolume totale è inferiore alla somma dei covolumi di tutte le molecole considerate una alla volta. Il risultato è che i tentativi di adattare l'equazione di van der Waals ai dati sperimentali conducono spesso a valori di *b* fisicamente inattendibili. Questa stima poco corretta della pressione a sfere rigide oscura il fatto che la teoria di van der Waals sia sostanzialmente corretta.

Negli anni cinquanta e sessanta numerosi ricercatori cominciarono a studiare l'applicabilità del modello a sfere rigide alla descrizione dei fluidi reali. Quindici anni fa abbiamo contribuito a tale ricerca formalizzando le idee di van der Waals sul modello di riferimento a sfere rigide in una teoria rigorosa, ma pratica. Nella nostra ricerca abbiamo impiegato idee avanzate da Alder, E. B. Smith, J. S. Rowlinson, Donald A. McQuarrie e Joseph L. Katz. Successivi miglioramenti sono stati introdotti da John D. Weeks, David Chandler e Hans C. Andersen.

Nel ricercare una formulazione rigorosa delle idee di van der Waals, il nostro metodo è quello di separare il potenziale intermolecolare in parti a breve e a lungo raggio d'azione. Il potenziale viene separato nel punto in cui esso si trasforma da attrattivo in repulsivo. Le due parti del potenziale servono in tal caso come base dei primi due termini di una serie di perturbazioni; la variabile elevata a potenze sempre più alte è il rapporto tra la profondità del potenziale e la temperatura. Il primo termine rappresenta la ripida repulsione a breve raggio d'azione, collegata al diametro della sfera rigida. Il secondo termine fornisce l'effetto predominante della parte attrattiva del potenziale, che è debole, ma a lungo raggio di azione.

Dal momento che il potenziale a breve raggio d'azione è molto ripido, è ragionevole supporre che il suo contributo all'energia libera sia virtualmente lo stesso che per la repulsione infinitamente ripida del modello a sfere rigide. Questa ipotesi fa parte di un passo cruciale nello sviluppo della nostra teoria. Abbiamo messo a punto un criterio matematico per specificare il diametro delle sfere di riferimento, un criterio che tiene conto delle interazioni repulsive in un liquido reale.

L'accurata scelta di un diametro per la sfera rigida è importante perché le molecole reali sono più molli delle sfere rigide e le loro funzioni energia potenziale sono «sfuggenti». In effetti il diametro di una molecola reale non è una grandezza fissa, ma dipende dall'energia con cui le molecole si urtano. Il diametro efficace deve perciò essere descritto da una media pesata della distribuzione repulsiva dell'energia potenziale, in modo tale che le proprietà termodinamiche del sistema a sfere rigide corrispondano alle proprietà mediate statisticamente di un sistema reale. Tale diametro efficace delle sfere rigide non è una costante, e non è neppure funzione della densità, ma dipende invece dalla temperatura, o dall'energia media delle molecole. Modificando il diametro, questo meccanismo consente al semplice, ma alquanto irreale, potenziale di sfere rigide di riprodurre il più dolce potenziale che governa il comportamento delle molecole reali.

Il secondo termine della nostra equazione di stato è un'espressione dell'energia totale del potenziale intermolecolare mediato su tutte le disposizioni molecolari del fluido di riferimento a sfere rigide. Questo termine corrisponde alla pressione interna dell'equazione di stato di van der Waals e si può rapidamente determinare dalla funzione di distribuzione radiale del fluido a sfere rigide. L'effetto del termine è principalmente quello di ridurre l'energia e la pressione del liquido.

Nessuno di tali termini riflette il fatto che le forze attrattive tra le molecole del tipo ipotizzato da van der Waals fanno differire la struttura del liquido da quella di un fluido di riferimento ideale della stessa densità. Le forze attrattive trascinano le molecole in regioni di minore densità e fanno perciò diminuire la loro energia libera. Alle alte densità l'effetto è piccolo a causa dello stretto addensamento delle molecole. Invece, a densità inferiori, la variazione di entropia diventa apprezzabile. Un terzo termine della nostra equazione spiega gli effetti a bassa densità e appare come una lieve differenza di entropia tra il fluido reale e quello di riferimento. Aggiungendo tali tre termini, la nostra somma di perturbazioni porta a risultati soddisfacenti per l'energia libera di un fluido. Il modello teorico è in eccellente accordo con gli esperimenti.

Van der Waals, nell'accettare il premio Nobel per la fisica nel 1910, commentò così l'inserimento della pressione interna nella sua teoria dei liquidi: «Confesso che si tratta di una strada tortuosa. Forse esiste una strada diretta. Nella ricerca di questa strada sarà una guida indispensabile *Elementary Principles in Statistical Mechanics* di Gibbs». La formulazione, negli anni sessanta, di una teoria simile ai risultati di van der Waals e i conseguenti chiarimenti e miglioramenti apportati dalla teoria delle perturbazioni e dalla meccanica statistica, possono essere interpretati come un completamento del programma iniziato da van der Waals negli anni settanta del secolo scorso.

L'epitelio dell'intestino tenue

È stato a lungo ritenuto un tessuto passivo; recenti ricerche hanno però dimostrato che le sue cellule possiedono una membrana in grado di demolire i cibi e di mettere in circolo le sostanze nutritive

di Florence Moog

Le sostanze nutritive entrano nell'organismo attraverso la bocca, ma, in realtà, esse diventano parte integrante dell'organismo solo dopo aver attraversato l'epitelio di rivestimento dell'intestino tenue ed essere passate nel sangue. Anche gli animali più evoluti, inclusi i mammiferi, possono essere descritti schematicamente come «un tubo entro un tubo». Il canale alimentare si estende dalla bocca all'ano, con la massa corporea situata tra questo stretto tubo e un altro più ampio costituito dalla pelle. Il compartimento interno del tratto gastrointestinale è quindi più strettamente collegato al mondo esterno che ai muscoli, ai nervi, allo scheletro e ai vasi sanguigni nell'organismo. Il passaggio di questa «frontiera» è l'evento fondamentale nel processo di assimilazione delle sostanze nutritive.

Una volta si pensava che la maggior parte dei processi digestivi avvenisse nel lume, o spazio interno, dell'intestino e che l'epitelio intestinale fosse una membrana passiva. Si riteneva anche che, dopo la digestione del cibo, le sostanze nutritive diffondessero semplicemente attraverso l'epitelio senza un intervento attivo da parte delle sue cellule. Ora, però, all'epitelio intestinale è stato attribuito un ruolo ben diverso. Recenti lavori nel campo della fisiologia intestinale hanno infatti dimostrato che la superficie interna dell'intestino è attiva sia nella demolizione sia nell'assorbimento delle sostanze nutritive. Le cellule di rivestimento possiedono enzimi legati alla membrana che convertono gli zuccheri complessi in zuccheri semplici e scindono i peptidi, provenienti dalla demolizione delle proteine, negli amminoacidi che li costituiscono o in piccoli peptidi, ciascuno formato da pochi amminoacidi.

Sembra inoltre che la membrana delle cellule dell'epitelio intestinale sia provvista di una varia gamma di sistemi di trasporto. Una volta completata la demolizione, questi sistemi di trasporto trasferiscono zuccheri, amminoacidi e piccoli peptidi nel citoplasma delle cellule da dove le sostanze nutritive vengono assorbite nel sangue. Sembra che per ogni semplice molecola di zucchero o amminoacido esistano parecchi meccanismi di trasporto, la maggior parte dei quali sono tuttora sconosciuti. La spiegazione di questi meccanismi è attualmente uno dei più interessanti filoni della ricerca nel campo della fisiologia dell'intestino. Non importa quale delle parecchie spiegazioni proposte sia alla fine la più soddisfacente: la teoria secondo la quale l'epitelio intestinale si comportava come una membrana passiva è stata comunque superata; si è infatti dimostrato che tale epitelio è un tessuto attivo che demolisce le sostanze nutritive, seleziona quelle che devono passare nel sangue e accelera il loro passaggio nell'organismo.

Dove e come avvengono la digestione e l'assorbimento sono problemi che hanno interessato a lungo i biologi. Le prime ricerche sono cominciate con esperimenti rozzi, ma chiarificatori, che coinvolgevano a volte lo stesso sperimentatore, fino ad arrivare a indagini più raffinate sulla struttura microscopica del rivestimento intestinale.

Nell'anno 1770 Lazzaro Spallanzani dell'Università di Pavia fece parecchi esperimenti che aiutarono a spiegare la natura del processo digestivo. Egli cominciò nutrendo il suo falco preferito con un pezzo di carne legato a una cordicella. Quando poco dopo estrasse la carne dallo stomaco dell'uccello notò che questa era più piccola essendo stata parzialmente digerita. Allora Spallanzani inghiottì alcuni pezzi di carne di pollo racchiusi in un sacchetto di stoffa. Quando riuscì ad espellere il sacchetto notò che parecchia carne non c'era più. In un esperimento leggermente più elaborato, Spallanzani pose nel tubo digerente del falco un tubo perforato contenente una spugna. Quando rimosse la spugna, ottenne da essa un liquido capace di digerire la carne.

Basandosi su questi esperimenti, Spallanzani stabilì che la digestione comincia nello stomaco a opera di sostanze capaci di attraversare la stoffa; sostanze che egli era riuscito in parte a estrarre con l'esperimento della spugna. Verso la metà del diciannovesimo secolo era noto che la digestione continua nell'intestino tenue dove tra gli agenti attivi sono inclusi enzimi secreti dal pancreas. Sebbene si fosse capito che l'intestino tenue partecipa alla digestione, poco si conosceva circa le modalità con cui avviene l'assorbimento delle sostanze nutritive.

Una spiegazione fu avanzata intorno al 1850 quando fu scoperto che molecole in soluzione possono diffondere attraverso una membrana semipermeabile, da una zona ad alta concentrazione verso un'altra avente concentrazione più bassa. Dopo un pasto le sostanze nutritive potrebbero perciò passare dal lume intestinale, avente concentrazione elevata, al sangue per diffusione passiva. Nei due decenni successivi però divenne chiaro che, sebbene la diffusione passiva potesse spiegare alcune delle proprietà assorbenti dell'epitelio intestinale, non poteva comunque giustificarle tutte. Intorno al 1870 era stato scoperto che è possibile osservare il passaggio di acqua da un segmento di intestino riempito con siero ematico al sangue. La diffusione passiva attraverso la parete intestinale non può spiegare questo fenomeno perché la concentrazione dell'acqua è approssimati-

La superficie interna dell'intestino tenue è rivestita da estroflessioni chiamate villi. In questa microfotografia elettronica a scansione, eseguita nel laboratorio dell'autrice presso la Washington University, i villi del duodeno, o tratto superiore dell'intestino tenue, di ratto sono ingranditi 630 volte. I villi conferiscono al tessuto intestinale un aspetto vellutato. Ogni villo è ricoperto da uno strato di epitelio prismatico semplice. Una volta si pensava che la digestione avvenisse esclusivamente nel lume, o spazio interno, dell'intestino. Recenti lavori nel campo della fisiologia hanno invece dimostrato che gran parte della digestione, come pure l'assorbimento, avviene a opera di strutture molecolari che sono legate alla membrana delle cellule epiteliali del villo.

vamente uguale nel sangue e nel siero. Inoltre era stato notato che le molecole di zucchero a sei atomi di carbonio utilizzate nei processi metabolici sono assorbite più rapidamente di quelle che non hanno importanza metabolica. Furono osservati anche fenomeni di assorbimento selettivo di grassi e proteine.

Gran parte della ricerca successiva sull'intestino tenue ebbe come scopo la scoperta dei mezzi attraverso i quali le molecole utili dal punto di vista fisiologico vengono demolite e quindi trasportate attraverso l'epitelio intestinale, velocemente e contro il gradiente di concentrazione. Simultaneamente veniva studiata la microanatomia di tale tessuto. Tutto ciò andò avanti fino a quando, intorno al 1960, questi due filoni della ricerca si sono incontrati grazie alla scoperta che alcune strutture molecolari della superficie interna dell'intestino sono importanti siti di digestione e assorbimento.

La struttura dell'intestino tenue è simile nella maggior parte dei mammiferi. Essa riflette due necessità: l'intestino deve fornire un'ampia superficie di assorbimento e deve adattarsi alla piccola cavità addominale. Queste condizioni sono state soddisfatte dall'avvolgimento dell'intestino e dallo sviluppo di parecchi tipi di estroflessioni sulla superficie interna dell'intestino che ne aumentano enormemente la superficie di assorbimento.

L'intestino tenue comincia dallo sfintere pilorico dello stomaco e termina all'inizio dell'intestino crasso. In un uomo adulto esso è lungo in media sei metri. L'intestino tenue può essere suddiviso in tre settori. Il duodeno discende per venticinque centimetri circa dallo stomaco e si attacca, per mezzo di un tessuto membranoso chiamato mesentere, alla parete posteriore della cavità addominale. La rimanente parte dell'intestino è facilmente movibile. Il tratto intestinale seguente al duodeno è chiamato digiuno e occupa la parte superiore sinistra della cavità addominale. L'ileo è situato più in basso nell'addome. In tutti e tre i tratti la parete mostra la stessa organizzazione stratificata. Tre strati esterni - la tonaca sierosa, la muscolare e la sottomucosa - circondano uno strato più interno chiamato mucosa, che è il più importante per la digestione e l'assorbimento delle sostanze nutritive. Nei mammiferi più grossi, incluso l'uomo, esistono tre tipi di estroflessioni che sporgono nel lume. In ordine di grandezza decrescente e di importanza fisiologica essi sono: le valvole conniventi, i villi e i microvilli. Le valvole conniventi o di Kerkring sono sporgenze laminari di otto-dieci millimetri di altezza che si estendono per metà fino a 2/3 della circonferenza interna dell'intestino. Queste strutture laminari sono presenti già a livello dell'estremità superiore del duodeno e scompaiono circa a metà dell'ileo. Esse sono più sviluppate nell'ultima parte del duodeno e nel segmento iniziale del digiuno. Le valvole conniventi possono triplicare la superficie intestinale.

Nell'intestino dei mammiferi più piccoli mancano le valvole conniventi; ci sono soltanto villi e microvilli. I villi si proiettano nel lume dalle pieghe e dalla circostante superficie della mucosa. Queste strutture, presenti con una densità di 10-40 villi per millimetro quadrato, conferiscono alla mucosa un aspetto vellutato, se guardata a occhio nudo. L'altezza dei villi varia da 0,5 a 1,5 millimetri. Tra i villi esistono depressioni chiamate cripte che affondano nella mucosa. All'interno di ogni villo c'è una rete di vasi sanguigni e linfatici sostenuti da un tessuto fibroso, lasso. È a questi due sistemi che devono essere trasportate le sostanze nutritive dopo la digestione.

Ogni villo è rivestito da uno strato di epitelio prismatico semplice in cui si distinguono due tipi principali di cellule: le cellule mucipare caliciformi che secernono il muco che ricopre e protegge la parete intestinale e le cellule assorbenti che esplicano le principali funzioni di digestione e assorbimento. Le cellule epiteliali che si trovano sul fondo della cripta formano una linea continua con quelle che rivestono il villo. Differiscono tuttavia da queste in quanto sono più piccole e più cubiche e hanno marcatori superficiali meno ben sviluppati. Inoltre in molte cellule della cripta, ma in nessuna dell'epitelio del villo, è possibile osservare figure mitotiche (cioè quelle configurazioni particolari che appaiono durante la divisione cellulare).

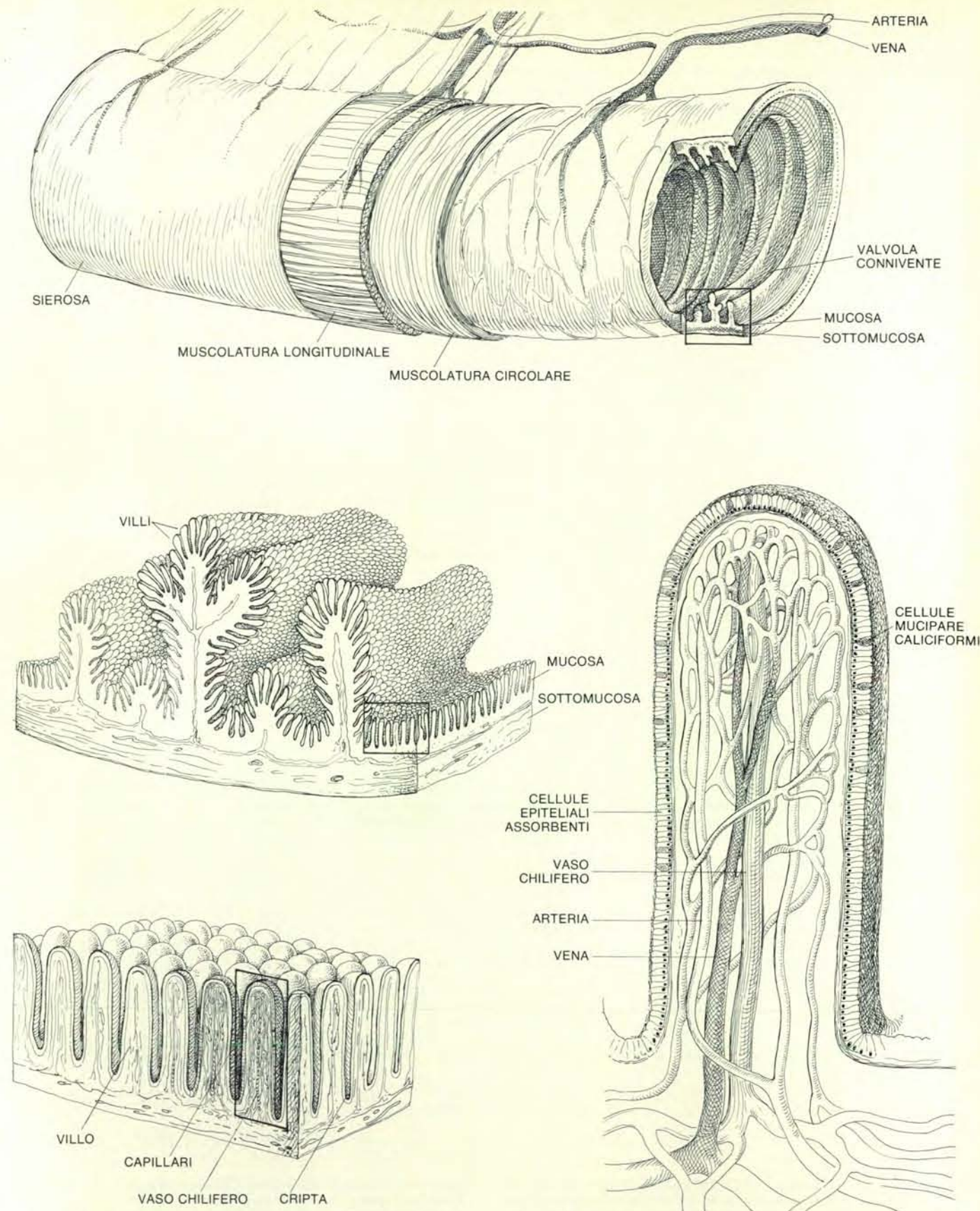
Osservazioni fatte con il microscopio ottico sulle differenze tra le cellule della cripta e quelle assorbenti dei villi suggeriscono che l'intestino può secernere un fluido digestivo (chiamato succo enterico), ma può anche assorbire le sostanze nutritive: le cellule della cripta sarebbero deputate al primo compito, le cellule assorbenti al secondo. Lavori successivi hanno dimostrato che, sebbene le cellule assorbenti assimilino le sostanze nutritive digerite, la relazione esistente tra i due tipi di cellule e il fluido digestivo è più sottile di quanto non si pensasse.

La chiave interpretativa di questa relazione è fornita dalle figure mitotiche scoperte nelle cellule della cripta. La distribuzione delle figure mitotiche indica una rapida proliferazione cellulare, insolita in una popolazione di cellule di un organismo adulto. Nel 1958 Charles P. Leblond e Bernard Messier della McGill University hanno dimostrato che le cellule della cripta migrano fino all'apice del villo. A questo scopo Leblond ha iniettato in alcuni ratti timidina marcata con trizio, l'isotopo radioattivo dell'idrogeno. La timidina è uno dei quattro nucleotidi che costituiscono le unità ripetitive della molecola di DNA; essa viene incorporata nel DNA solo quando la cellula sta per dividersi. La timidina è l'unico dei quattro nucleotidi del DNA a non essere anche un componente dell'RNA; il suo alto contenuto nelle cellule è quindi un segno evidente della replicazione del DNA. Poche ore dopo l'iniezione, le cellule marcate con l'isotopo radioattivo sono state trovate solo nella cripta. Ventiquattr'ore dopo, la timidina radioattiva è stata trovata in cellule situate a metà strada sul villo; dopo quarantotto ore le cellule radioattive sono state trovate sulla sommità del villo.

Il ciclo di proliferazione e migrazione delle cellule epiteliali del villo richiede più o meno lo stesso periodo di tempo in tutti i mammiferi. Dopo che queste cellule hanno raggiunto la sommità del villo, si staccano e cadono nel lume, portando con sé tutti gli enzimi digestivi che possiedono. Quindi gli enzimi trovati nel succo enterico non sono secreti, nel senso comune del termine, ma sono immessi, insieme al rimanente contenuto delle cellule assorbenti, nel lume intestinale. Circa diciassette miliardi di cellule vengono liberate giornalmente lungo tutto l'intestino umano.

La spiegazione del ciclo migratorio tra le cellule dell'epitelio ha stimolato l'interesse nei confronti di queste cellule. Ulteriori studi hanno dimostrato che le cellule assorbenti e le cellule caliciformi si differenziano in quanto migrano lungo il villo. In questo processo ogni tipo cellulare acquisisce le sue strutture caratteristiche.

Una cellula assorbente completamente sviluppata è alta e stretta. Il nucleo è situato profondamente nel citoplasma, sotto un'ampia zona apicale riccamente fornita di mitocondri, reticolo endoplasmatico e vescicole di Golgi; l'abbondanza di questi organelli indica un'intensa attività di ossidazione e sintesi. Proprio sotto la faccia superiore, o apice, si nota che ogni cellula assorbente è legata alle cellule vicine da speciali dispositivi di giunzione presenti sulla membrana cellulare. Le molecole non possono penetrare attraverso queste giunzioni occludenti, e così le sostanze nutritive non possono entrare nei vasi sanguigni, presenti all'interno del villo, attraverso canali intercellulari, come si pensava una volta. Sotto queste giunzioni vi sono i desmosomi, che formano connessioni meno compatte tra le membrane cellulari. Le giunzioni occludenti e i desmosomi fanno sì che l'epitelio che ricopre il villo formi un foglietto continuo.

La caratteristica che distingue le cellule assorbenti è un orletto striato di 1,5 micrometri, all'apice della cellula, che si affaccia nel lume intestinale. La struttura dell'orletto supera di poco il potere di risoluzione del microscopio ottico, e questo ha messo in difficoltà gli istologi fino al 1950, quando Barbara Granger e Richard F. Baker dell'Università della Southern California sono riusciti a dimostrare, mediante il microscopio elettronico, che il bordo è composto da un terzo tipo di estroflessioni. Queste estroflessioni più piccole, i microvilli, sono lunghe circa un micrometro e la loro densità è di circa 200 000 per millimetro quadrato nel digiuno. La loro presenza nell'orletto a spazzola, così ora è chiamato l'orlo striato, accresce di venti volte la superficie intestinale. Mentre in un primo tempo si pensava che quest'area fosse di circa 15 metri quadrati, ora si sa che essa è di circa 300 metri quadrati.

La struttura dell'intestino tenue è simile in tutte le specie di mammiferi. La parete intestinale è composta da quattro strati concentrici. I tre strati esterni - la tonaca sierosa, la muscolare (costituita da due strati muscolari) e la sottomucosa - circondano la mucosa interna, la struttura più importante per la digestione e l'assorbimento. Parecchi tipi di estroflessioni che sporgono dalla superficie intestinale aumentano la sua area assorbente. Nei mammiferi più grossi, incluso l'uomo, le estroflessioni più grandi sono le valvole conniventi. Alte 8-10 millimetri, esse ricoprono da metà a 2/3 della superficie intestinale interna (sono più sviluppate nel digiuno, il segmento centrale dell'intestino). I villi sporgono nello spazio interno a partire dalle pieghe e dalla superficie circostante. Nell'intestino umano l'altezza dei villi varia da 0,5 a 1,5 millimetri; la loro densità varia da 10-40 villi per millimetro quadrato. Dentro ogni villo c'è una fitta rete di vasi sanguigni e linfatici ai quali arrivano, a digestione completa, le sostanze nutritive. Tra i villi vi sono depressioni chiamate cripte che si spingono fino alla mucosa.

All'interno di ogni microvillo c'è un nucleo di fibre che sostiene l'organello. In ogni cellula assorbente tali fibre partono da tutti i microvilli e terminano in uno strato di fibre chiamato tessuto basale, così compatto che può essere paragonato a un feltro. Il microvillo è ricoperto dalla membrana cellulare formata da due strati contrapposti di molecole lipidiche. Ogni molecola lipidica comprende una «testa» idrofila e due «code» idrofobe costituite da lunghe catene di acidi grassi. Nella membrana le teste polari si affacciano all'esterno e le code, nei due strati, sono disposte una di fronte all'altra. La membrana del microvillo ha un rivestimento di carboidrati che è più spesso di quello della maggior parte delle membrane cellulari.

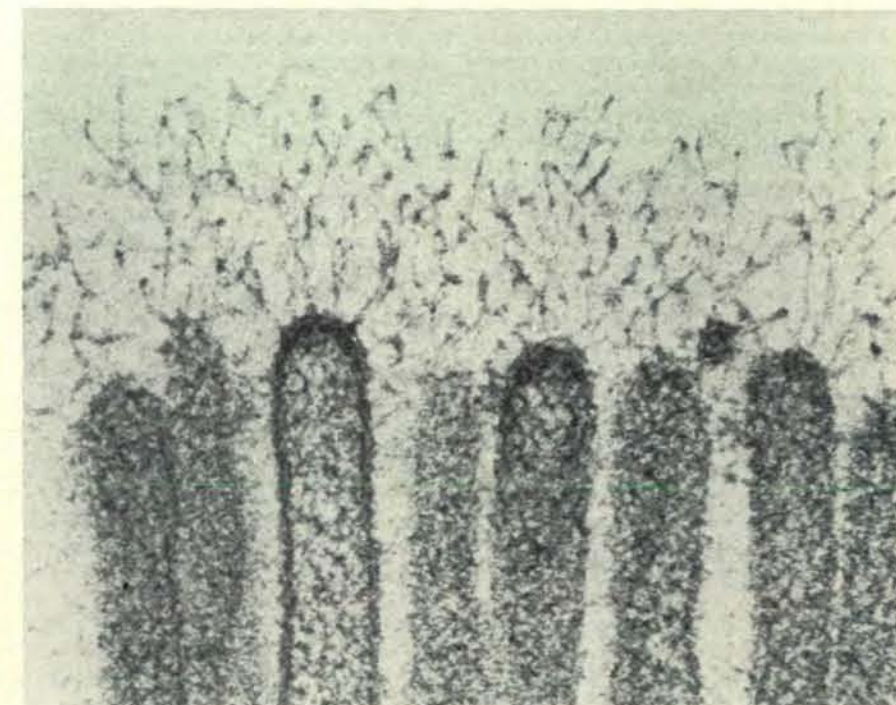
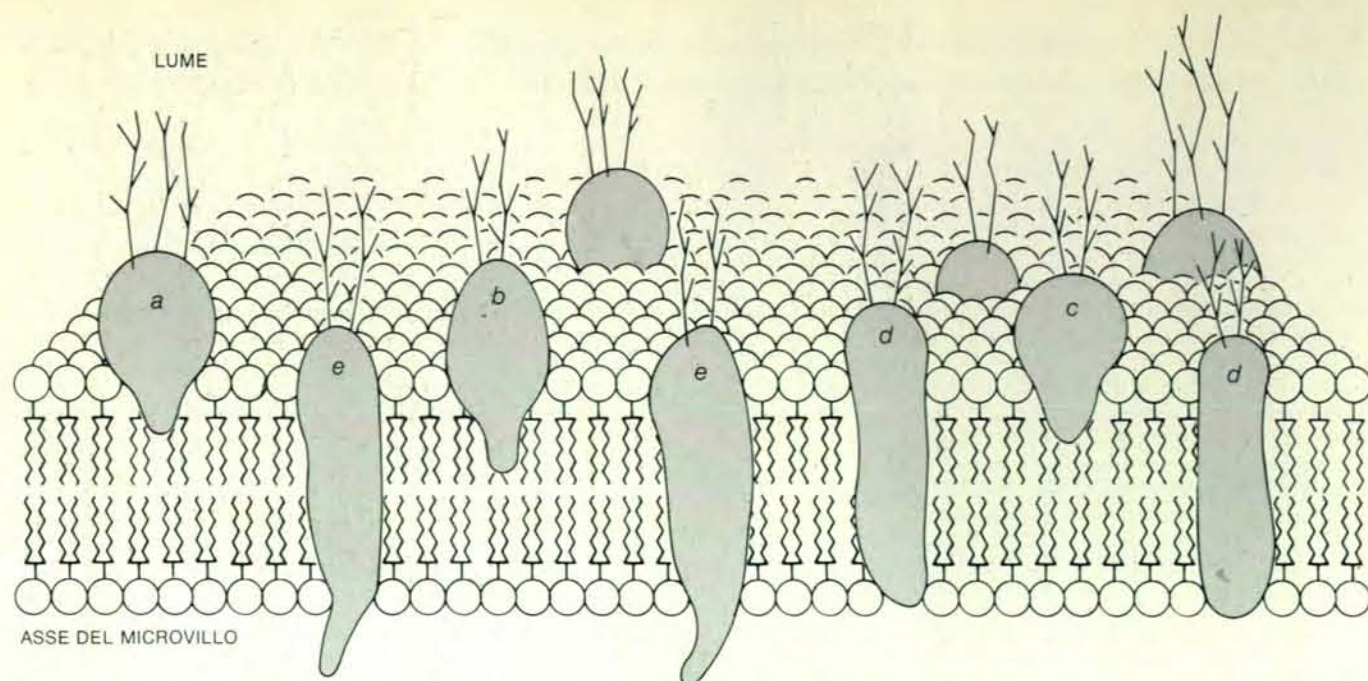
È proprio la membrana che ricopre il microvillo la zona attraverso la quale le sostanze nutritive entrano effettivamente nell'organismo. Una serie di ricerche svolte in molti laboratori ha dimostrato che nella membrana è incorporato un complesso di enzimi e sistemi di trasporto. Gli enzimi sono in grado di demolire una vasta gamma di sostanze alimentari, riducendole in molecole che sono poi trasportate nel citoplasma delle cellule epiteliali dai sistemi di trasporto. Alle volte le molecole vengono ulteriormente trasformate nel citoplasma delle cellule assorbenti prima di essere trasportate nei vasi sanguigni all'interno del villo.

La struttura delle cellule assorbenti dell'epitelio intestinale riflette la loro specifica funzione nella digestione. La cellula assorbente costituisce uno dei due tipi principali di cellule che formano l'epitelio del villo. L'altro tipo è costituito dalle cellule caliciformi che secernono muco. Entrambi i tipi di cellule derivano dalla divisione di cellule presenti nelle cripte; assumono poi la loro struttura specifica durante una migrazione verso l'apice del villo che richiede due giorni. Qualche frammento della membrana di copertura delle cellule assorbenti può staccarsi poiché, durante questo tragitto, le cellule sono esposte all'azione dei fluidi digestivi. Le cellule sono ricche di mitocondri, che forniscono energia per i processi metabolici; di reticolo endoplasmatico dove vengono sintetizzate le proteine (inclusi gli enzimi digestivi); di vescicole di Golgi, dove alle proteine vengono attaccate catene laterali di carboidrati. Le giunzioni occludenti e i desmosomi collegano tra loro le cellule assorbenti formando un unico strato che ricopre il villo. Il nucleo si trova in profondità, nella cellula. Il carattere più specifico della cellula è un orletto a spazzola, una struttura presente sulla superficie luminale della cellula. L'orletto a spazzola è formato da tante file di piccole estroflessioni chiamate microvilli. All'interno del microvillo c'è un fascio di fibre costituite da proteine. In ogni cellula, le fibre che partono dai microvilli raggiungono il tessuto basale, una struttura a forma di rete in cui le fibre corrono parallele alla superficie luminale della cellula.

Susumu Ito della Harvard University è stato il primo a stabilire che il rivestimento più esterno della membrana ricco di carboidrati, conosciuto come glicocalice, è composto da fibre sottili che si continuano nella membrana stessa. Negli anni quaranta, George Gomori dell'Università di Chicago e H. Takematsu dell'Università di Tokyo svilupparono una tecnica in grado di dimostrare che in una cellula un enzima può occupare una posizione fissa. Essi usarono questo metodo per dimostrare che l'enzima fosfatasi alcalina, presente nelle cellule assorbenti, è localizzato a livello dell'orletto a spazzola. La fosfatasi alcalina idrolizza, o rompe, una varietà di composti fosfati, molti dei quali presenti nell'alimentazione dell'uomo. Negli anni sessanta l'analisi chimica dei tessuti combinata con la microscopia elettronica ha dimostrato che la presenza della fosfatasi alcalina è limitata alla membrana dei microvilli. La fosfatasi alcalina e tutti gli altri enzimi legati alla membrana dell'orletto a spazzola sono glicoproteine, ossia proteine con catene laterali formate da carboidrati.

Nel 1961 David Miller e Robert K. Crane della Washington University di St. Louis misero a punto una tecnica per isolare un lembo intatto di orletto a spazzola da omogenati di mucosa intestinale. Isolando l'orletto in tal modo Miller e Crane furono in grado di analizzarlo chimicamente. Essi trovarono tra i suoi costituenti gli enzimi saccarasi e maltasi, che sono disaccaridasi, ossia enzimi che rompono le molecole a 12 atomi di carbonio in subunità a 6 atomi di carbonio. La saccarasi scinde una molecola di saccarosio in una molecola di glucosio e una di fruttosio. La maltasi scinde il maltosio (un derivato dell'amido) in due molecole di glucosio. Da allora, nell'orletto a spazzola, sono state scoperte la lattasi (che scinde il lattosio, o zucchero del latte, in una molecola di glucosio e una di galattosio) e parecchie altre disaccaridasi. Indagini in altri laboratori hanno rivelato, nell'orletto a spazzola, la presenza anche di amminopeptidasi, cioè di enzimi che rimuovono l'amminoacido finale dall'estremità amminica di una corta catena peptidica.

Durante gli anni settanta fu accertato che questi enzimi fanno parte della membrana del microvillo, anziché essere collegati all'orletto a spazzola in altro modo. Ciò è stato reso possibile mediante l'uso di nuove tecniche che consentono di separare la membrana del microvillo dal rimanente orletto a spazzola. Al microscopio elettronico la frazione di membrana isolata dalle cellule assorbenti appare come un insieme di vescicole chiuse con estroflessioni nodose sulla superficie esterna. Da questi preparati può essere estratta una frazione che ha la stessa attività degli enzimi digestivi.

In questa microfotografia elettronica è visibile, ingrandito 18 000 volte, l'orletto a spazzola di una cellula assorbente del duodeno di topo. La struttura dell'orletto a spazzola è appena al di sopra del potere di risoluzione del microscopio ottico. Fu nel 1950 che, mediante il microscopio elettronico, si dimostrò come l'orletto a spazzola fosse costituito da singole estroflessioni, i microvilli. Questi sono lunghi circa un micrometro e hanno una densità di 200 000 per millimetro quadrato nel digiuno umano. Essi aumentano di circa 20 volte la superficie dell'intestino tenue. La membrana esterna di ogni microvillo è rivestita dal glicocalice, un rivestimento polisaccaridico più spesso di quello della maggior parte delle membrane cellulari. Qui il glicocalice appare come una tenue masserella che si allunga verso l'alto dall'apice dei microvilli. La linea scura sulla destra rappresenta il confine tra due cellule assorbenti. La microfotografia è stata eseguita nel laboratorio dell'autrice.

Il glicocalice si allunga dall'apice del microvillo nel lume intestinale. Questa microfotografia elettronica, eseguita negli anni sessanta da Susumu Ito della Harvard University, mostra i microvilli dell'intestino ileo di un gatto, ingranditi 140 500 volte. Essa ha dimostrato per la prima volta in modo convincente che le fibre del glicocalice sono inserite nella membrana del microvillo. Ora si sa che alcune fibre sono parti di molecole di enzimi digestivi incorporate nella membrana. Gli enzimi sono glicoproteine ossia proteine a cui sono legate parecchie catene di carboidrati; la parte glicosidica si protende nel lume. Agli enzimi della membrana dei microvilli è affidato gran parte del lavoro di demolizione degli zuccheri, delle proteine e dei composti a base di fosforo.

Gli enzimi digestivi glicoproteici penetrano a varie profondità nella membrana dei microvilli. La componente proteica dell'enzima è localizzata a livello della zona lipidica della membrana; le catene glicosidiche sporgono nel lume. I lipidi della membrana sono disposti in due strati, con le «teste» idrofile rivolte verso l'esterno e le «code» idrofobe, composte da catene di acidi grassi, affacciate all'interno. Gli enzimi inseriti nella membrana sono di diversi tipi e si differenziano a seconda delle sostanze che demoliscono. Ogni disaccaridasi *(a, b, c)* scinde un solo tipo di zucchero a dodici atomi di carbonio in due subunità a sei atomi di carbonio. Questi enzimi sporgono dalla membrana nella quale penetrano solamente fino a una profondità limitata. La fosfatasi alcalina *(d)* idrolizza, o scinde, la maggior parte dei fosfocomposti presenti nel cibo; la componente proteica dell'enzima è la misura della profondità della membrana. Le amminopeptidasi *(e)* rimuovono un amminoacido da un'estremità di una corta catena peptidica; passano poi completamente attraverso la membrana e all'interno dei microvilli.

Dalle vescicole della membrana si possono estrarre facilmente disaccaridasi attive mediante l'uso di papaina, un enzima proteico che scinde i legami che uniscono gli enzimi alla membrana. Invece l'estrazione della fosfatasi alcalina e delle amminopeptidasi dalle vescicole richiede l'impiego di un detergente capace di rompere i legami tra le molecole lipidiche. Queste scoperte fanno ritenere che le disaccaridasi sono senz'altro legate solamente alla parte esterna della membrana; infatti, una proteasi come la papaina può staccarle da essa. La fosfatasi alcalina e le amminopeptidasi sono invece situate più profondamente e la loro estrazione richiede la rottura dei legami tra le molecole lipidiche situate dentro la membrana.

S. Maroux e i suoi collaboratori dell'Istituto di biochimica e di biologia molecolare di Marsiglia hanno dimostrato che le amminopeptidasi non soltanto penetrano nella membrana, ma sporgono anche dalla sua superficie interna, verso la parte interna del microvillo. La più convincente tra le attuali teorie circa le relazioni tra la membrana e gli enzimi digestivi sostiene che la componente proteica della glicoproteina penetra a varie profondità, in relazione all'enzima; la componente polisaccaridica si estende al di fuori, nel lume, formando il glicocalice che, nelle microfotografie elettroniche, appare come una rete di sottili fibre.

Gli enzimi glicoproteici compiono gran parte del lavoro di digestione direttamente sulla superficie del microvillo. L'attività enzimatica della membrana del microvillo, tuttavia, varia considerevolmente lungo il villo. Una volta si pensava che gli enzimi legati alla membrana e i sistemi di trasporto fossero sintetizzati mentre le cellule assorbenti erano ancora nelle cripte ed entrassero in funzione quando l'epitelio si spostava verso l'apice del villo. Ora è chiaro che la sintesi proteica nelle cellule assorbenti avviene senza intenzione mentre l'epitelio si sposta verso la sommità del villo.

Milton M. Weiser, che lavora nel laboratorio di Kurt J. Isselbacher ad Harvard, ha messo a punto una tecnica che permette di riunire le cellule situate esternamente sul villo in serie ordinata, dall'apice alla cripta. La procedura richiede il riempimento di un tratto di intestino con una soluzione tampone seguita da una soluzione contenente una sostanza chiamata acido etilendiamminotetracetico (EDTA). Dopo un periodo di circa un'ora, le cellule fluttuano libere, cominciando da quelle situate all'apice per finire con quelle della cripta. Con questo metodo Weiser ha separato le cellule epiteliali in nove frazioni, confermando i risultati precedenti secondo cui l'attività della fosfatasi alcalina e della saccarasi è assente nelle cellule situate nel profondo della cripta, ma cresce costantemente procedendo verso l'apice del villo.

Weiser ha usato la sua tecnica per spiegare i processi che stanno alla base di questo gradiente nell'attività enzimatica. Egli ha dimostrato che molecole di zucchero a sei atomi di carbonio sono incorporate nelle glicoproteine (incluse la fosfatasi alcalina e le disaccaridasi) più rapidamente nelle cellule della parte superiore del villo che non in quelle della base. È noto che molecole di zucchero marcate sono assunte dalle cellule assorbenti e penetrano nelle vescicole di Golgi entro cinque minuti dall'iniezione (nei ratti). Le vescicole di Golgi costituiscono un sistema di membrane nelle quali i carboidrati vengono uniti alle proteine neosintetizzate, per formare una glicoproteina. Quando questo processo si è completato, la nuova molecola di glicoproteina è trasportata a livello dell'orletto a spazzola.

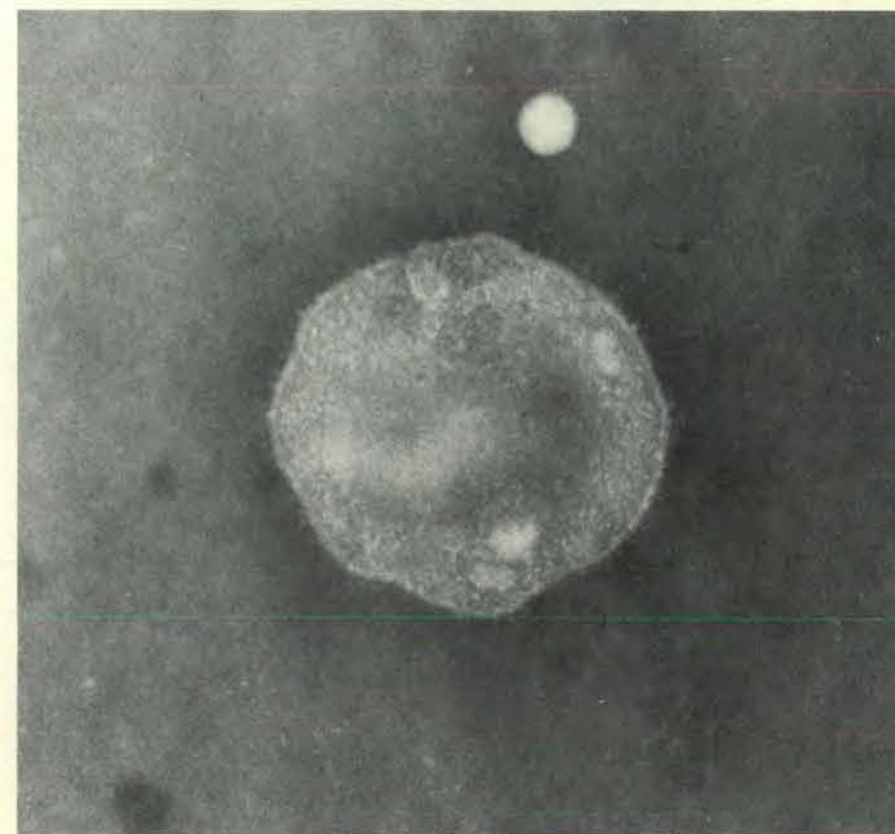
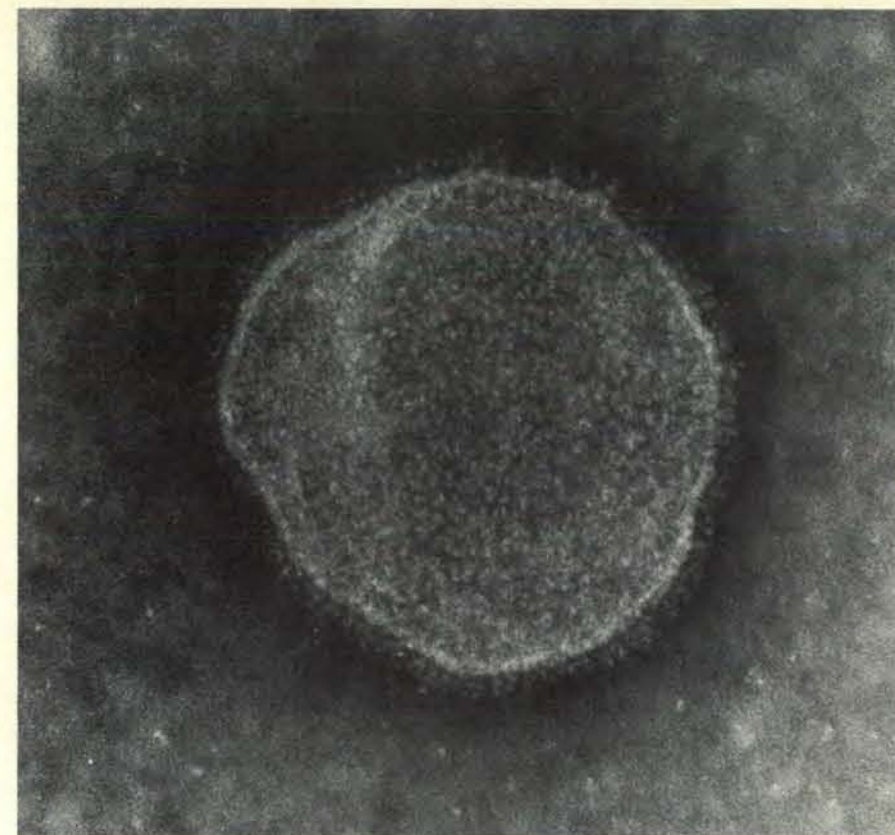
Weiser ha calcolato la percentuale di aumento dello zucchero mettendo le cellule estratte dal villo in una soluzione di glucosammina marcata con un isotopo radioattivo del carbonio: la glucosammina è uno zucchero amminico incorporato nelle catene laterali di carboidrati di alcune glicoproteine. Le cellule provenienti dalla metà superiore del villo assumono il marcatore radioattivo due volte più velocemente delle cellule situate più sotto. Le cellule della cripta incorporano solo piccole quantità di zucchero marcato.

Weiser ha confermato questi risultati iniettando in ratti una soluzione di molecole di zucchero marcate e quindi impiegando EDTA per raccogliere frazioni di cellule del villo in ordine sequenziale. Di nuovo, le cellule situate all'apice del villo hanno incorporato più marcatore di quanto abbiano fatto le cellule vicino alla base.

Il fatto che le cellule del villo producano RNA sostiene la teoria di una sintesi proteica più abbondante nella zona apicale del villo. Un isotopo radioattivo del fosforo iniettato nei ratti o nei pulcini è assorbito dalle cellule dell'epitelio intestinale, che lo incorporano nelle molecole di RNA in formazione. Le cellule della cripta incorporano fosforo radioattivo sia nell'RNA ribosomico sia nell'RNA messaggero. (L'RNA ribosomico costituisce una parte delle strutture necessarie per la costruzione di tutte le proteine; l'RNA messaggero specifica la sequenza di amminoacidi in ogni singola molecola proteica.) Lungo il villo invece il fosforo radioattivo è incorporato solamente in un tipo di RNA messaggero, chiamato RNA poliadenilato. (Ciò è stato confermato nel 1980 da Alan Morrison e John W. Porteus dell'Università di Aberdeen.) Una specifica molecola di RNA poliadenilato è necessaria per la sintesi di ogni specie di proteina. Le molecole di RNA poliadenilato sono instabili e quindi devono essere fabbricate ininterrottamente per assicurare la quantità richiesta di glicoproteine indispensabili. D'altra parte l'RNA dei ribosomi è stabile e non deve essere rinnovato durante la migrazione delle cellule verso l'apice del villo.

Perché viene impiegata tanta energia per la sintesi proteica e per il montaggio delle glicoproteine nelle cellule assorbenti quando queste hanno vita breve ed evidentemente sono così ben differenziate? La risposta risiede nel fatto che l'orletto a spazzola esposto al movimento del chimo, cioè del cibo parzialmente digerito, e agli enzimi digestivi presenti nel lume, è distrutto rapidamente e deve essere continuamente riparato non appena migra verso l'alto. David H. Alpers della Washington University Medical School ha dimostrato che i legami che uniscono alla membrana le disaccaridasi situate sulla superficie sono soggetti all'attacco delle proteasi secrete dal pancreas. Una disaccaridasi sopravvive sulla membrana solo poche ore. Inoltre, man mano che la digestione procede la membrana può perdere alcuni frammenti. Interi microvilli possono rompersi. Soltanto una continua sintesi può mantenere la struttura necessaria per la digestione e l'assorbimento.

Una volta che i cibi sono stati ridotti ai loro componenti molecolari dagli enzimi presenti nel lume intestinale e dagli enzimi glicoproteici annessi alla membrana dei microvilli, le sostanze nutritive devono essere immesse nel torrente circolatorio; in questo modo si rendono disponibili a tutti i tessuti. Sebbene siano stati compiuti molti progressi circa la relazione esistente tra gli enzimi glicoproteici e la membrana dei microvilli, rimangono abbastanza sconosciuti i meccanismi per mezzo dei quali le sostanze nutritive vengono trasportate attraverso la membrana.

L'amido, la sorgente più importante di carboidrati nella maggior parte dei cibi che consumiamo, è ridotto a maltosio dall'amilasi, nella bocca e nel lume intestina-

Un tratto di membrana di un orletto a spazzola isolato si presenta come un insieme di vescicole chiuse con protuberanze costituite da glicocalici, sulla superficie esterna. Nella microfotografia elettronica, in alto, le vescicole sono ingrandite 263 000 volte. Quando le vescicole sono trattate con papaina, un enzima proteolitico, le protuberanze scompaiono, come è mostrato nella microfotografia in basso, in cui l'ingrandimento è di circa 237 000 diametri. Le disaccaridasi attive possono essere liberate dalle vescicole usando la papaina, che spezza i legami che uniscono le disaccaridasi alla superficie esterna della membrana. La fosfatasi alcalina e le amminopeptidasi possono essere liberate solo da una sostanza capace di rompere la matrice lipidica della membrana. Le microfotografie elettroniche sono state eseguite da D. Maestracci dell'Università di Montreal.

le; l'amilasi è un enzima secreto dalle ghiandole salivari e dal pancreas. Alcune molecole di amilasi possono essere imprigionate nelle fibre del glicocalice insieme ad altri enzimi pancreatici. Come già detto, la maltasi legata alla membrana dell'orletto a spazzola scinde ogni molecola di maltosio in due molecole di glucosio.

Quando per la prima volta fu dimostrato che gli enzimi glicoproteici sono legati alla membrana, sembrò logico che una molecola di maltasi, dopo aver scisso il disaccaride, dovesse servire come molecola di trasporto (o *carrier*) per gli zuccheri a sei atomi di carbonio verso l'interno del microvillo. Tale meccanismo può spiegare alcuni tipi di trasporto di molecole di glucosio nelle cellule epiteliali assorbenti. Ma ora si sa tuttavia che la maltasi della membrana può scindere molto più maltosio di quanto questo sistema di trasporto possa trasportare; alcune delle molecole di glucosio liberate vengono immesse nel lume.

Dopo aver circolato nel lume, il glucosio può ritornare sulla superficie del villo. Si pensa che esista un secondo sistema di trasporto che fa attraversare la membrana alle molecole di glucosio. Questo secondo sistema sarebbe in grado di pompare le molecole di glucosio nelle cellule assorbenti. In questo sistema il movimento delle molecole di glucosio è accoppiato al trasporto di ioni sodio attraverso la membrana nella stessa direzione; questo tipo di movimento è stato osservato anche in altre cellule. Il flusso parallelo di ioni sodio è in grado di fornire l'energia necessaria per l'accumulo del glucosio in concentrazione superiore a quella presente nel lume. Si presume che un difetto in questo sistema di trasporto del glucosio sia il responsabile di una condizione chiamata «malassorbimento o assorbimento difettoso del glucosio-galattosio», caratterizzata da grave diarrea. Sebbene si abbia un'idea di quello che potrebbe essere un secondo meccanismo di trasporto, non si conosce ancora l'esatta via molecolare attraverso la quale le molecole di zucchero entrano nelle cellule assorbenti.

La natura della via attraverso la quale gli amminoacidi e i piccoli peptidi (cioè corte catene amminoacidiche) entrano nelle cellule assorbenti è ancora più oscura. Mentre i carboidrati arrivano alla membrana dei microvilli come semplici coppie di zuccheri a sei atomi di carbonio, le proteine vengono parzialmente digerite nel lume, formando così un insieme di singoli amminoacidi e di piccole catene peptidiche. Negli anni cinquanta si pensava che le proteine fossero completamente digerite dalle proteasi secrete dallo stomaco e dal pancreas, lasciando solamente singoli amminoacidi. In realtà dopo la digestione nel lume intestinale restano non solo amminoacidi liberi, ma anche catene peptidiche di varie lunghezze. Verso la fine degli anni sessanta si accumularono prove che dimostravano l'assorbimento diretto dei dipeptidi (molecole costituite da due amminoacidi) nelle cellule epiteliali. Recentemente il quadro relativo all'assorbimento delle proteine si è ulteriormente complicato. Ora è chiaro che anche i tripeptidi, costituiti da tre amminoacidi, possono penetrare direttamente nelle cellule assorbenti.

Sono stati proposti parecchi meccanismi per spiegare come le molecole di zucchero sono trasportate, attraverso la membrana del microvillo, nel citoplasma della cellula assorbente. La disaccaridasi che scinde uno zucchero a dodici atomi di carbonio, come il saccarosio, nei suoi componenti a sei atomi di carbonio, è legata alla superficie della membrana. Se l'orientamento dell'enzima rimane invariato durante il processo litico, i due zuccheri possono essere liberati nello spazio tra i microvilli *(a)*. Potrebbero quindi essere trasportati attraverso la membrana da un meccanismo indipendente dall'enzima glicoproteico. Il sito attivo dell'enzima, che trattiene la molecola di zucchero, può essere ruotato verso l'interno mentre la molecola di zucchero viene spezzata. In questo caso gli zuccheri possono essere liberati all'interno della membrana stessa *(b)*, oppure dentro un canale acquoso che si estende attraverso la membrana nel citoplasma della cellula assorbente *(c)*.

Sembra che per l'assorbimento degli amminoacidi siano necessari quattro sistemi attivi di trasporto, chiaramente indipendenti l'uno dall'altro, nonostante la loro natura molecolare non sia ancora conosciuta. Questi sistemi corrispondono alle quattro classi nelle quali i biochimici hanno suddiviso i venti o più amminoacidi con i quali l'intestino potrebbe venire a contatto. La classe più numerosa, i quindici amminoacidi non carichi o neutri, hanno il gruppo amminico ($-NH_2$) e carbossilico ($-COOH$) legati allo stesso atomo di carbonio. Questa configurazione è necessaria perché l'amminoacido possa essere riconosciuto dalla molecola trasportatrice. Gli amminoacidi non carichi sono trasportati più rapidamente, ma a velocità che variano in rapporto all'affinità dell'amminoacido per la molecola trasportatrice; le molecole non cariche competono tra loro per entrare nella cellula assorbente. Anche i tre amminoacidi bibasici, aventi due amminogruppi, competono tra loro per un comune sito di trasporto. Esiste una piccola competizione incrociata tra gli amminoacidi neutri e quelli bibasici per entrare nelle cellule assorbenti. La glicina, l'amminoacido più semplice, la prolina e l'idrossiprolina, nei quali l'atomo di azoto del gruppo amminico fa parte di una struttura ad anello, hanno affinità per una terza molecola trasportatrice. Anche l'acido glutammico e l'acido aspartico, aventi due gruppi carbossilici, hanno un loro specifico sistema. È difficile stabilire se questo quarto sistema di trasporto porti gli amminoacidi dentro la cellula contro il gradiente di concentrazione oppure no, perché questi amminoacidi sono immediatamente trasformati in glutammina e asparagina. Il trasporto di amminoacidi di tutti i quattro sistemi richiede energia ed è accoppiato a un flusso verso l'interno di ioni sodio.

L'esistenza dei quattro sistemi di trasporto degli amminoacidi è dimostrata da alcune rare malattie dovute a un errato trasporto di un solo tipo di amminoacido. Nella malattia di Hartnup, per esempio, l'intestino non è in grado di assorbire alcuni amminoacidi neutri, mentre altri vengono assorbiti senza difficoltà.

Uno specifico meccanismo di trasporto distinto da quelli per i singoli amminoacidi provvede al passaggio di dipeptidi e tripeptidi attraverso la membrana del microvillo. Quando un tetrapeptide arriva a livello dell'orletto a spazzola può subire a opera di una tetrapeptidasi la rimozione dell'amminoacido finale: rimane così un tripeptide e un amminoacido libero. Il tripeptide può essere assorbito direttamente nel citoplasma, oppure può essere scisso, da un secondo enzima, in un dipeptide e in un amminoacido libero. Gli enzimi che scindono i tripeptidi sono distribuiti abbastanza uniformemente tra la membrana e il citoplasma, quelli che dividono i dideptidi sono per lo più all'interno della cellula. Pochi dipeptidi sono scissi sulla membrana; anzi, la maggior parte di essi sono trasportati intatti all'interno della cellula assorbente.

Il trasporto di peptidi piuttosto che amminoacidi presenta parecchi vantaggi per l'organismo. Non si ha competizione tra gli amminoacidi per i siti nei quattro sistemi di trasporto delle specifiche molecole quando peptidi interi attraversano lo strato lipidico della membrana. La presenza di sistemi di trasporto peptidici riduce al minimo l'energia impiegata per muoversi contro il gradiente di concentrazione; dipeptidi e tripeptidi sono velocemente scissi all'interno della cellula, cosicché la concentrazione delle catene peptidiche non raggiunge un livello più alto di quello del lume. Inoltre un difetto in uno dei quattro sistemi di trasporto dei singoli amminoacidi può essere supplito dal trasporto peptidico, dal momento che le molecole che sarebbero dovute entrare nella cellula tramite il sistema di trasporto difettoso, entrano come componenti di catene peptidiche. La struttura molecolare dei sistemi di trasporto per i peptidi, come quella per i singoli amminoacidi, non è stata ancora chiarita. Esistono ulteriori lacune nella conoscenza della struttura interna del microvillo. Fino agli inizi degli anni settanta si pensava che il microvillo fosse una struttura rigida, che serviva esclusivamente per incrementare la superficie della membrana. Nel 1975 Mark Mooseker, che allora lavorava presso l'Università della Pennsylvania, dimostrò che un frammento di orletto a spazzola isolato poteva contrarsi in determinate condizioni ambientali. Tra le condizioni indispensabili è inclusa l'esposizione del preparato agli ioni calcio e all'adenosintrifosfato (ATP), la molecola che partecipa a una serie di reazioni a catena che liberano l'energia necessaria per la maggior parte dei processi cellulari.

Quando un microvillo si contrae, le fibre interne scorrono verso la membrana sottostante. Mooseker e Lewis G. Tilney, anch'egli dell'Università della Pennsylvania, avevano dimostrato che l'orletto a spazzola contiene actina e miosina, le proteine che scorrono l'una sull'altra durante la contrazione del muscolo. Nel tessuto muscolare una terza proteina, la meromiosina, media l'interazione tra l'actina e la miosina. Mooseker e Tilney misero la meromiosina a contatto con il microvillo per determinare se anche in questo caso esiste un'interazione simile a quella osservata nel muscolo. Avendola osservata proposero che durante la contrazione del microvillo i filamenti di actina, all'interno, si spostassero vicino a quelli orizzontali della miosina nel tessuto basale. Essi suggeriscono che la meromiosina faciliti la contrazione del microvillo consentendo all'actina di agganciarsi ai filamenti di miosina.

Nel tentativo di spiegare come e perché il microvillo si contrae, sono state isolate altre quattro proteine. La più interessante di queste è la calmodulina, una proteina simile alla troponina, che lega gli ioni calcio e inoltre regola la contrazione scheletro-muscolare. È noto che la calmodulina ha funzioni analoghe in cellule non muscolari. Essa potrebbe essere d'aiuto nel regolare la contrazione del microvillo, ma la sua presenza può essere spiegata in altro modo. La calmodulina è concentrata più a livello dello stelo del microvillo piuttosto che nel tessuto basale, dove interagiscono actina e miosina. Poiché il calcio è una delle sostanze che deve essere assorbita a livello dell'epitelio intestinale, Mooseker ha suggerito che la calmodulina potrebbe essere un costituente del sistema per il trasporto degli ioni calcio. Nel trasportare gli ioni la proteina potrebbe migrare tra la parte interna del microvillo e il tessuto basale. Se questa ipotesi fosse corretta, potremmo capire i sistemi di trasporto sia nella membrana che nel citoplasma delle cellule assorbenti.

Gli amminoacidi e i corti peptidi, cioè i prodotti della demolizione delle proteine, sono trasportati attraverso la membrana dei microvilli da una serie di sistemi di trasporto. Gli enzimi proteolitici, nel lume intestinale, scindono le proteine in singoli amminoacidi e catene peptidiche. La membrana dei microvilli ha quattro sistemi per trasportare i singoli amminoacidi nel citoplasma delle cellule assorbenti. Tutti gli amminoacidi aventi simile struttura chimica competono per entrare nella cellula utilizzando lo stesso sistema di trasporto. Il movimento degli amminoacidi verso l'interno della cellula è accoppiato a un flusso parallelo di ioni sodio, un processo che fornisce l'energia necessaria per l'accumulo di amminoacidi a una concentrazione superiore a quella presente nel lume intestinale. Le amminopeptidasi legate alla membrana staccano dalle corte catene peptidiche l'amminoacido finale. Dopo aver subito questo accorciamento, la catena peptidica può ritornare nel lume intestinale, oppure può attraversare direttamente la membrana. Per il trasporto di dipeptidi e tripeptidi esiste un meccanismo specifico. Dopo aver attraversato la membrana, i piccoli peptidi sono ridotti a singoli amminoacidi dall'azione di peptidasi citoplasmatiche. Sebbene le linee generali di questo processo, chiarito tra il 1960 e il 1970, siano ormai confermate, la struttura molecolare delle molecole trasportatrici, in molti casi, non è ancora nota.

Una stazione basca per la caccia alla balena nel Labrador

Meno di cent'anni dopo la scoperta dell'America, in acque canadesi c'era una fiorente industria spagnola di caccia alla balena. Per la prima volta ne sono state trovate prove in archivi, sulle coste e sul fondo marino

di James A. Tuck e Robert Grenier

La costa orientale del Canada è stata zona di pesca per gli europei almeno a partire dal 1497, quando Giovanni Caboto completò il primo viaggio al golfo del San Lorenzo di cui si abbia notizia. Caboto riferì di avervi trovato merluzzi in tale abbondanza che li si poteva pescare calando semplicemente in mare una cesta dal bordo della nave. Le acque canadesi dell'Atlantico divennero ben presto una mecca per i pescatori europei, in particolare per quelli originari del golfo di Biscaglia.

Nel Cinquecento arrivarono in tali acque marinai biscaglini, in particolare baschi, non solo per pescarvi il merluzzo, ma anche per praticarvi la caccia alla balena. Essi uscivano in mare da stazioni sulla costa, dove si insediavano per mesi, e dove fondevano il grasso di balena, facendo poi ritorno in patria col loro prezioso carico d'olio di balena all'inizio dell'inverno. I baschi esercitavano la loro caccia annuale in una località chiamata da loro Grand Bay. Fino a poco tempo fa, però, erano venute in luce ben poche prove di quest'industria, per quanto prospera.

All'inizio degli anni settanta ricerche condotte in archivi spagnoli cominciarono a fornire informazioni sull'attività di caccia alla balena nella Grand Bay. Successive ricerche archeologiche, stimolate dai documenti cinquecenteschi, hanno rivelato i resti di una delle stazioni costiere basche sulla costa meridionale del Labrador. Esse hanno condotto anche alla scoperta, nelle acque costiere in prossimità della stazione, del relitto di una nave basca, che conteneva oggetti notevolmente ben conservati. In questo articolo diremo qualcosa sulle ricerche d'archivio (ancora in corso) che hanno dato origine alle scoperte archeologiche e riferiremo su ciò che è stato trovato nel sito del Labrador negli scavi condotti negli ultimi quattro anni. Gli scavi sulla costa sono stati condotti da ricercatori della Memorial University di Terranova; il lavoro sul fondo marino è stato compiuto dalla Marine Excavation Unit of Parks Canada (l'equivalente canadese dell'US National Park Service). A partire dallo scorso anno, quello che si è trasformato in un progetto multidisciplinare, che potrebbe richiedere per essere completato altre tre stagioni di lavoro, ha ricevuto aiuti dai Parks Canada, dal Canadian Social Sciences and Humanities Research Council, dal governo di Terranova e Labrador, dagli Archivi di stato del Canada e dal Canadian Conservation Institute of the National Museums of Canada.

Le ricerche d'archivio in Spagna vengono condotte da Selma Huxley Barkham, una studiosa della storia dei baschi. Il suo progetto di lavoro ha richiesto la traduzione e l'analisi di migliaia di documenti notarili, come contratti, noli, cause giudiziarie, cambiali e testamenti che contengono riferimenti all'attività della pesca nella Grand Bay. La signora Barkham ha trovato questi documenti, molti dei quali erano da parecchio tempo dimenticati, in vari archivi baschi, in particolare a Oñate. Fra gli indizi da lei scoperti vi sono nomi di luoghi baschi per indicare varie località nell'area della Grand Bay. I documenti conservano notizia anche di numerosi casi di buona o cattiva fortuna nella caccia alla balena. Per esempio, la ricognizione eseguita dalla Marine Excavation Unit, che trovò il relitto di nave, fu intrapresa sulla base di una relazione del 1565 sulla perdita di una nave, la *San Juan*, poco prima del previsto ritorno in Spagna con un carico di olio di balena.

Gli studi della Barkham hanno già fornito ampie prove della produttività della caccia alla balena in acque canadesi nel Cinquecento. Un documento attesta che nel 1571 Doña Marina de Urançu assicurò un quinto di un carico d'olio di balena del Nuovo Mondo per 2000 ducati. Questo fatto suggerisce che il valore dell'intero carico fosse di 10 000 ducati. Tale somma risulta essere press'a poco uguale a quella spesa per l'acquisto di due grandi galeoni a Siviglia nello stesso anno. Perciò in una singola stagione fortunata di caccia alla balena, un armatore (o un gruppo di armatori) poteva recuperare il costo di una nuova nave e procurarsi anche un profitto consistente.

I documenti ci forniscono anche almeno alcune informazioni essenziali sull'attività di caccia alla balena. Le navi partivano di solito dirette alla Grand Bay in primavera. Un documento di nolo del 1566 menziona come data di partenza di una nave il 25 aprile, ma di solito le navi partivano più tardi. Una volta arrivati alla loro stazione sulla costa canadese, i balenieri uscivano a caccia con scialuppe, imbarcazioni lunghe di solito circa cinque metri. Essi continuavano la caccia finché le condizioni del ghiaccio non li costringevano a tornare in Spagna in dicembre o forse addirittura in gennaio. Spesso accadeva che dovessero partire in tutta fretta a causa dell'improvvisa formazione di ghiacci marini (la banchisa) o di ghiacci costieri. Qualche volta la fuga dai ghiacci era impossibile e i balenieri dovevano trascorrere l'inverno nella stazione. Particolarmente rigidi furono gli anni attorno al 1575; le ricerche della Barkham hanno dimostrato che i balenieri svernarono alla Grand Bay, contro la loro volontà, sia nell'inverno 1574-1575 sia nel 1576-1577.

Nell'inverno del 1576-1577 i balenieri baschi ebbero numerose perdite. Alcuni decessi furono dovuti forse al freddo o alla mancanza di provviste, ma almeno uno, avvenuto nel giugno 1577, quando l'inverno era passato da un pezzo, potrebbe essere attribuibile allo scorbuto o a qualche altra malattia da carenza. Come apprendiamo dalle ultime volontà testamentarie della vittima (è questo il più antico documento noto di questo genere che a quanto sappiamo sia stato redatto in Canada), il defunto si chiamava Juan Martínez de Larrume. Nonostante i rischi associati alla caccia alla balena con piccole imbarcazioni, la maggior parte dei de-

Vista da nord del porto interno di Red Bay, nel Labrador, qui in primo piano. Sullo sfondo si scorge lo stretto di Belle Isle, oltre la Saddle Island e la Penney's Island, più piccola. Qui, nel Cinquecento, una colonia di balenieri baschi fondò una stazione costiera di caccia alla balena.

La sezione centrale di un relitto, probabilmente la nave *San Juan*, naufragata nel 1565, fu localizzata da archeologi marini nel 1978 al largo di Saddle Island. La fotografia illustra il paramezzale, tre costole, la coperta del lato di tribordo, una griglia di riferimento, zavorra e barili.

cessi documentati alla Grand Bay sono connessi ai soggiorni invernali forzati più che a incidenti di caccia.

Le ricerche della Barkham hanno fornito molte altre informazioni sulle tecniche della caccia alla balena a partire da stazioni costiere e sulla vita dei balenieri. La preda preferita era la balena boreale o *Balaena mysticetus*, che si ciba di krill, piccoli crostacei planctonici; era nota agli inglesi come la balena «giusta» (forse perché era la balena giusta da cacciare; nuotava infatti lentamente e, dopo essere stata colpita a morte, galleggiava anziché colare a picco e andare in tal modo perduta). La tecnica di caccia è la stessa che viene descritta con tanta vivacità in *Moby Dick:* la balena veniva colpita con arpioni e, dopo che si era indebolita quanto bastava, veniva uccisa con una lancia. (Nel 1566 il capitano Miguel de Cerain imbarcò sulla sua nave, diretta alla Grand Bay, 100 arpioni e 24 lance; gli arpioni andavano perduti più facilmente delle lance.) La balena veniva poi rimorchiata sino alla stazione costiera, dove uomini armati di grossi coltelli ne staccavano a strisce il grasso. Quando l'animale doveva essere rimorchiato controcorrente e contro vento, il lavoro doveva essere duro; un documento attesta che a volte una balena veniva ormeggiata in attesa di condizioni favorevoli. Nel 1575 Nicholas de la Torre fu accusato di essersi appropriato del grasso di una balena che era stata uccisa dall'equipaggio di Joan López de Rey, che l'aveva poi lasciata ormeggiata a uno scoglio al largo della costa. Il processo che ne seguì durò per vent'anni; quando si concluse, sopravvivevano solo le vedove dei due capitani.

Di particolare interesse per le informazioni che ci forniscono indirettamente sulla vita quotidiana nelle stazioni costiere sono i dati sulle provviste imbarcate su navi dirette verso la Gran Bay. La Barkham ha tradotto finora due elenchi di provviste. Fra i cibi c'erano pancetta affumicata, merluzzo salato (anche se il merluzzo poteva essere pescato nelle acque del Nuovo Mondo), sardine, farina, fave, piselli, olio d'oliva, semi di senape, aglio, vino di Xeres, sidro e grandi quantità di gallette. L'equipaggiamento per la caccia alla balena comprendeva - oltre alle scialuppe, agli arpioni e alle lance - funi di canapa di varia grandezza e coltelli per staccare e tagliare il grasso. Sono elencate anche, ma non specificate, medicine per il barbiere della nave, che era anche un cerusico in quanto si occupava delle operazioni di chirurgia spicciola.

Fra i materiali da costruzione portati in Canada c'erano chiodi, «terra» (presumibilmente argilla) e tegole. Non viene menzionato il legname, cosicché pare di poter congetturare a colpo sicuro che gli edifici delle stazioni costiere dovevano essere costruiti con legname locale; le asce portate dalla Spagna dovevano servire al taglio di alberi sul posto. (Il capitano de Cerain, partendo per la Grand Bay, imbarcò fra le altre cose 16 asce.) In che modo questi edifici venissero costruiti viene in gran parte ignorato nei documenti. Erano richieste tegole per tetti; nel 1563 Francisco de Florriaga imbarcò una «quantità di tegole sufficiente per la riparazione delle capanne» rimaste esposte alle intemperie l'inverno precedente. Un solo accenno, per di più indiretto, ci suggerisce in che modo venissero usate tali capanne. Testimoniando in un processo, uno degli uomini addetti al lancio degli arpioni, Simón de Azcoita, riferiva di aver tenuto una conversazione con un compagno «accanto alla capanna che avevano costruito per bollire le balene che uccidevano».

Dov'era la Grand Bay? Studiando i nomi di località citati nei documenti, la Barkham è giunta alla conclusione che i balenieri baschi dovevano aver dato quel nome allo stretto di Belle Isle, l'angusto braccio di mare fra la parte più settentrionale di Terranova e la costa meridionale del Quebec-Labrador, che immette dall'Atlantico nel golfo del San Lorenzo. Un luogo menzionato nei documenti baschi era Samadet o Semadet, nome che sembra potersi identificare con East St. Modeste, un antico insediamento pochi chilometri a ovest della Red Bay sulla riva labradoriana dello stretto. Un altro nome di luogo basco era Xateau o Chateo, che potrebbe corrispondere all'attuale Chateau, un altro insediamento sulla costa del Labrador. Il gran numero di riferimenti a un terzo luogo, variamente chiamato Butus, Buytes, Boytus e Buitres, suggerisce che dovesse trattarsi di una fra le stazioni costiere basche più importanti.

La Red Bay è uno fra i migliori porti naturali sulla costa del Labrador in prossimità dello stretto. Essa comprende un bacino interno circondato da colline molto ripide, alte da 90 a 150 metri. È riparata rispetto alla baia esterna da una bassa isola, con un piccolo rilievo a ciascuna estremità, che i cartografi dell'Ammiragliato britannico hanno chiamato, non senza ragione, Saddle Island (Isola a sella). Sembrava opportuno compiere una visita alla Red Bay.

Perciò nel 1977 la Barkham e uno di noi (Tuck), accompagnati da Walter Kenyon, un rappresentante del Royal Ontario Museum, e da altre persone, visitarono la cittadina di Red Bay. Trovammo le spiagge e gli orti della città cosparsi di migliaia di frammenti di tegole rosse. La stessa scoperta fu fatta sulle coste della Saddle Island, dove vari muri di pietra recavano tracce di una sostanza nera solidificata. Alcuni campioni della sostanza furono in seguito identificati come residui di grasso animale bruciato. Ne concludemmo che i frammenti di tegole dovevano essere i resti delle tegole usate dai baschi nella costruzione dei tetti. Il materiale nero sui muri di pietre suggeriva che era qui che il grasso di balena veniva fuso in olio.

Le nostre scoperte erano abbastanza promettenti da consentirci di procurarci un finanziamento per un mese di scavi esplorativi l'anno seguente. In quella prima stagione di scavi il lavoro fu concentrato in tre aree sulla Saddle Island che avevano l'aria di essere rimaste praticamente indisturbate dopo essere state abbandonate 400 anni prima. Su una piccola terrazza che dominava il porto scoprimmo i resti di un tetto di tegole rosse caduto. Qui doveva dunque essere esistito un edificio; ai suoi resti erano associati ceramiche, oggetti in vetro, in ferro e altri manufatti, che consentirono di datare l'edificio a un'epoca imprecisata della seconda metà del Cinquecento.

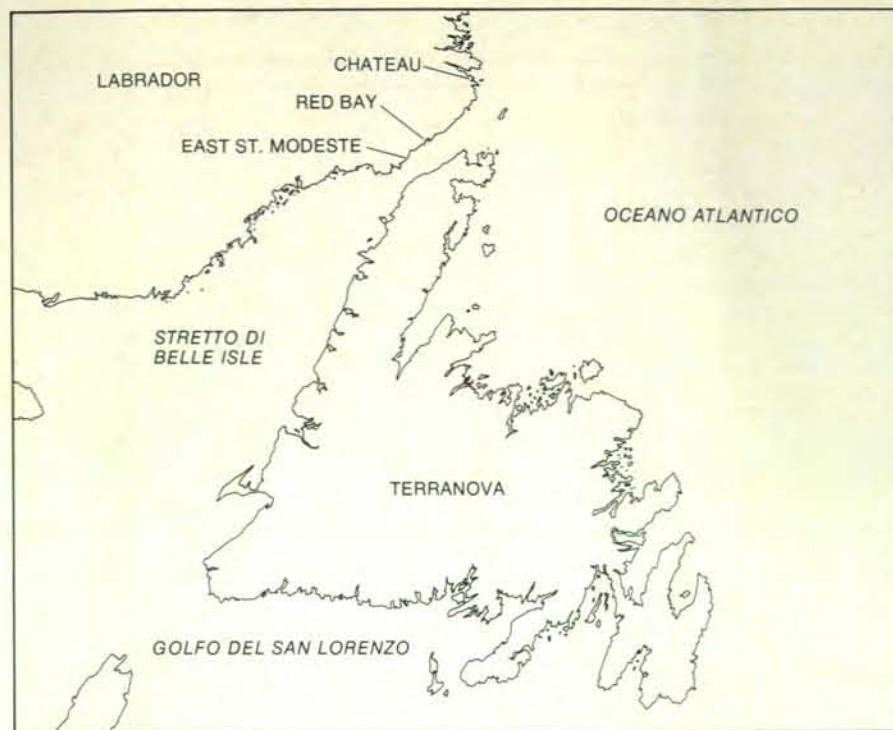
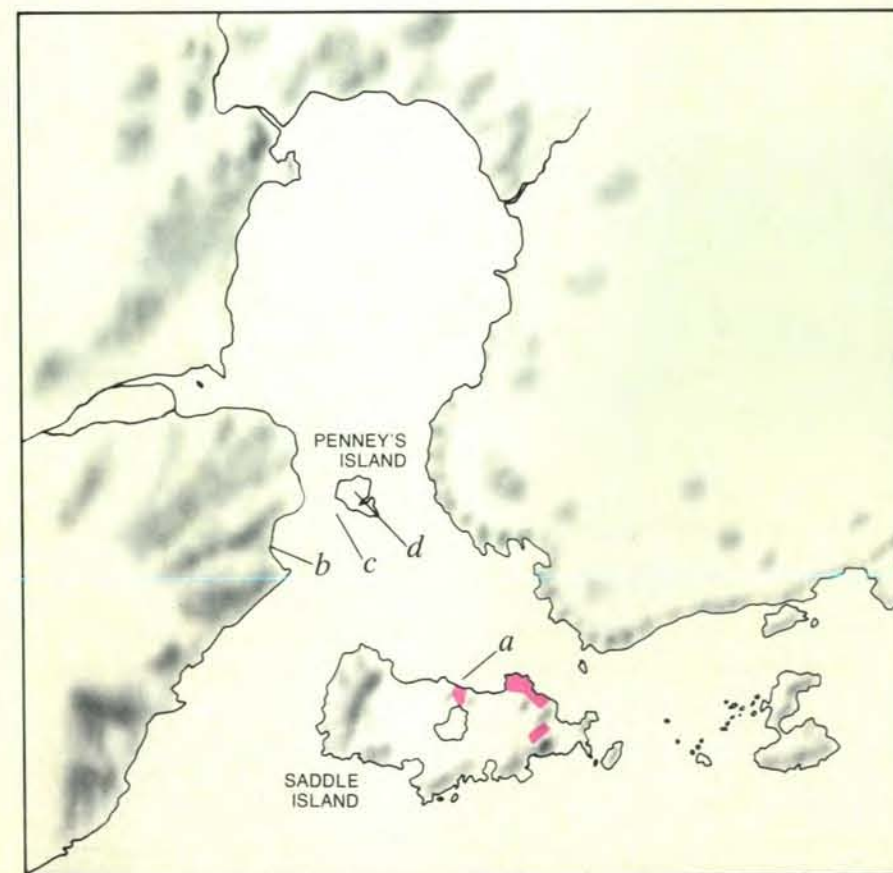
Il secondo scavo esplorativo ebbe luogo in un'area imbevuta d'acqua, il tipo di luogo in cui è più probabile la conservazione di resti organici. La trincea di prova da noi scavata riportò in luce staia di schegge di legno, il risultato della lavorazione di tronchi d'albero abbattuti sul luogo per ridurli alla forma e alle dimensioni volute. Furono scoperti anche parti di cerchi e legacci di barili, frammenti di cuoio e alcuni pezzetti di quercia, un albero che non è originario del Labrador.

Il terzo sito prescelto era situato dove erano stati trovati i residui di grasso animale, un luogo che già ritenevamo fosse stato sede dell'attività di fusione del grasso, ovvero un'area su cui dovevano sorgere dei forni. Una piccola trincea di prova rivelò uno strato di tegole ben conservato sepolto sotto uno strato di sabbia rimescolato dalle onde. Il muro di pietre che aveva attratto la nostra attenzione in occasione della nostra prima visita si trovava vicino al centro di quelli che erano chiaramente i resti di una grande struttura.

Il gruppo della Memorial University aveva già lasciato Red Bay, più che soddisfatto per i risultati della stagione di scavi, quando i subacquei del Parks Canada fecero la più spettacolare scoperta singola del 1978. A soli 30 metri al largo della costa settentrionale di Saddle Island, essi localizzarono i resti di una nave affondata con un carico di barili di legno. I cerchi e i legacci dei barili corrispondevano esattamente a quelli da noi trovati nel sito imbevuto d'acqua. Inoltre la nave concordava nei particolari con la descrizione di un naufragio registrato negli archivi baschi. Era stata la conoscenza del naufragio della *San Juan* da parte della Barkham a condurre alla ricognizione sottomarina dell'unità di subacquei dei Parks Canada.

Nella stagione seguente, oltre a compiere ricerche in nuove zone di Saddle Island, gli archeologi disegnarono una cartina dell'area in cui nel 1978 era stato scoperto il tetto caduto, designata come Area *A*, la fotografarono e ne rimossero i frammenti di tegole. La zona delineata dalle tegole era lunga 14 metri e larga 8, cosicché la struttura ricoperta dal tetto potrebbe essere stata solo leggermente più piccola. Evidentemente essa era stata costruita almeno in parte con pietre trovate sul posto. Altri muri in pietra sull'isola hanno tipicamente un nucleo composto da ghiaia e piccole pietre e una faccia formata da pietre di dimensioni maggiori, disposte con cura l'una sull'altra. Qui, però, solo una linea di pietrisco suggeriva dove doveva trovarsi il muro frontale, quello rivolto a nord, della struttura. Non ci fu bisogno di costruire un muro posteriore perché la struttura era ubicata in

Un tetto caduto, fatto di tegole rosse, viene esaminato da archeologi della Memorial University di Terranova. Fra le rovine, in prossimità della costa settentrionale di Saddle Island, furono trovati numerosi resti di manufatti prodotti in Spagna nella seconda metà del Cinquecento.

Ossa di balena imbiancate coprono un tratto di spiaggia lungo la parte occidentale del porto di Red Bay. Gli archeologi hanno trovato qui i resti di almeno 30 balene boreali (della specie *Balaena mysticetus*) e altre ossa di balena sono state rinvenute al largo di Penney's Island.

Lo stretto di Belle Isle, l'ingresso al golfo del San Lorenzo, che separa Terranova dalla costa del Quebec-Labrador, era noto ai balenieri baschi del Cinquecento come Grand Bay. East St. Modeste potrebbe identificarsi con una stazione costiera dei baschi da loro chiamata Semadet, Chateau con la stazione denominata Xateau, mentre Red Bay corrisponderebbe alla stazione di Butus.

Qui è ingrandita l'area di Red Bay. La maggior parte degli scavi sono stati compiuti nelle parti di Saddle Island segnate in colore; la nave affondata si trova in *a*. Il deposito di ossa di pagina 103 si trova in *b*, sulla costa. Altre ossa di balena sono state rinvenute al largo della Penney's Island, così come il relitto di una nave più piccola (*c*). Forni per la fusione del grasso sono stati trovati in *d*.

modo tale che come parete sud veniva utilizzato un affioramento roccioso quasi verticale. (L'incorporazione di tali elementi naturali nelle costruzioni è un uso basco comune.)

I muri più corti a est e a ovest, se sono mai esistiti, non hanno lasciato traccia. Può darsi che quest'edificio fosse aperto alle due estremità, o forse le due pareti corte erano in legno e imputridirono rapidamente nel suolo acido dell'isola. È chiaro che l'intelaiatura del tetto era in legno, come attestano centinaia di chiodi fucinati a mano. Non possiamo però stabilire se fosse un tetto a uno o due spioventi. Nelle costruzioni basche entrambi gli stili sono comuni.

In ogni caso quest'edificio ospitò due piccoli focolari che furono riportati in luce durante gli scavi della stagione 1980. È risultato che essi contenevano carbone di legna, ossa di animali bruciate, frammenti di ceramica, pallini di piombo e frammenti di scarti della fusione del piombo. Era evidente che, oltre a fornire calore, i focolari servivano per cucinare e per altri compiti domestici.

Nell'edificio e attorno ad esso (specialmente all'estremità occidentale, dove c'era forse una discarica di rifiuti) sono stati trovati molti manufatti. I più numerosi erano frammenti di ceramica, per un totale di più di 2000. Molti di essi poterono essere identificati come parti di giare per la conservazione di cibi, vasellame di cucina, brocche, scodelle e piatti. I cocci di una parte delle brocche, delle tazze e delle anfore di dimensioni minori recavano decorazioni azzurre, verdi o dorate applicate su una vernice bianca. Tali oggetti in maiolica (nome derivato dall'isola di Maiorca, che era uno fra i principali centri di produzione) erano d'uso comune nella Spagna del Cinquecento (dalla quale venivano anche esportati). Molto meno abbondanti erano gli oggetti in vetro; sono stati infatti riportati in luce solo i frammenti di un calice da vino e di una grande bottiglia simile a una caraffa, decorata con un disegno in rilievo di forcelle di pollo.

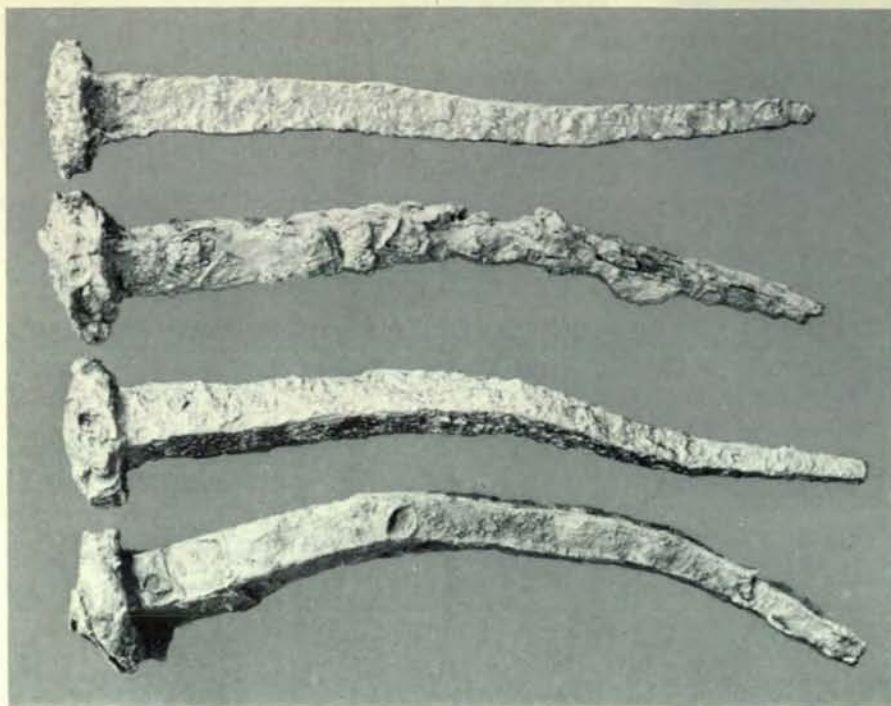
Oltre ai chiodi in ferro usati per fissare il telaio in legno dell'edificio, gli scavatori trovarono un certo numero di altri manufatti in metallo. Fra questi c'erano vari coltelli, uno dei quali aveva un manico di legno intagliato in modo elaborato e un pomo in bronzo in forma di corona. Congiuntamente ai focolari e al carattere domestico della maiolica, i coltelli suggerirono che quest'edificio fosse usato come abitazione, impressione confermata da una varietà di altri reperti: un rosario in legno, una moneta e una chiave. Al termine della stagione di scavi del 1979 sembrava plausibile che avessero vissuto qui individui di stato sociale relativamente elevato.

I risultati ottenuti la stagione successiva in una località adiacente, sempre sulla Saddle Island, l'Area *E*, modificarono quest'opinione. Una struttura riportata in luce in quest'area si trovava in un sito quasi identico a quello dell'Area *A*, e la maggior parte dei manufatti qui rinvenuti - fra cui ceramiche, oggetti in vetro e altri oggetti di uso domestico - erano molto simili a quelli dell'Area *A*. Altri oggetti indicavano però che gli occupanti di questa abitazione erano bottai. Fra questi utensili vi sono due accette, una delle quali è così ben conservata che il suo bordo tagliente è ancora affilato, una mazza per montare i cerchi e varie morse per sistemare il fondo superiore del barile quando venivano montate le doghe. Questi reperti gettarono nuova luce su due pezzetti di metallo che erano stati trovati in precedenza nell'Area *A* e che non furono identificabili finché il Canadian Conservation Institute non ne ebbe fatto lastre ai raggi X. Uno di essi risultò essere la lama di una pialla e l'altro un frammento di una lama di sega. Pare perciò probabile che i residenti di entrambe le aree fossero impegnati in attività di lavorazione del legname.

Ogni dubbio su questo fatto fu ben presto fugato, per quanto concerne l'Area *E*. Immediatamente a nord di questa struttura, in depressioni naturali del substrato roccioso, si era accumulata dell'acqua freatica. Quando compimmo scavi in queste depressioni sature d'acqua, trovammo che contenevano scarti dell'attività di costruzione di barili così ben conservati che sembrava vi fossero stati gettati solo da qualche anno e non da qualche secolo. C'erano parti di barili che erano state scartate a causa di qualche imperfezione nel materiale o di qualche difetto di lavorazione. Doghe e pezzi di fondi di quercia e di faggio corrispondevano a quelli che erano stati recuperati nel relitto e lo stesso vale per i cerchi fatti di ontano spaccato e per i legacci di vimine dei barili. C'erano frammenti di legno residui dalla lavorazione dei pezzi arrotondati dei fondi, estremità di cerchi tagliate via, frammenti di salice troppo nodosi o troppo piccoli per poter essere usati, e anche trucioli e segatura. Senza dubbio l'edificio dell'Area *E* combinava le funzioni di abitazione e di laboratorio per la costruzione di barili e la somiglianza generale riscontrata fra esso e la struttura dell'Area *A* suggerisce che entrambe le costruzioni dovevano svolgere lo stesso ruolo. Nondimeno la nostra valutazione del 1979 non doveva essere completamente sbagliata; forse i costruttori di barili, la cui attività era certamente essenziale per poter trasportare in patria l'olio di balena, godevano nelle stazioni costiere dei balenieri baschi, di uno stato sociale elevato.

Fra l'alloggiamento dei bottai e la costa c'erano i focolai per la fusione del grasso di balena. Uno di essi, che aveva un muro in pietra che si è conservato per l'altezza di un metro, fu riportato in luce durante la stagione del 1980. Il muro, che correva per una decina di metri parallelo alla linea di costa, formava la «spina dorsale» dell'edificio dei forni. Dalla parte verso costa c'erano le grandi caldaie per la fusione del grasso, e dalla parte verso l'interno c'era la zona di lavorazione. Il muro era costruito con cura con pietre locali e con le pietre usate come zavorra dalle navi nel

Fondi e doghe di barile recuperati dal carico della nave affondata davanti a Saddle Island si sono conservati perfettamente nei quattro secoli in cui sono rimasti immersi nelle fredde acque del Labrador. Sono ancora visibili i segni dei cerchi di rami di ontano che legavano assieme le doghe.

Questa piccola brocca, che è uno dei numerosi oggetti di maiolica tipicamente spagnola rinvenuti in quest'area, fu trovata nel sito di una capanna in cui lavoravano i bottai di una stazione costiera.

Questi grossi chiodi, lunghi 17-18 centimetri, erano fra le centinaia di chiodi fucinati a mano usati dai balenieri baschi nella costruzione delle loro case e dei capannoni dei forni su Saddle Island.

viaggio di andata, cementate con una malta fatta con argilla (probabilmente la «terra» elencata fra i materiali da costruzione importati). L'argilla, impregnata con olio di balena, legava le pietre con una presa simile a quella del cemento.

Come avveniva la fusione del grasso di balena? Non possediamo alcuna descrizione delle tecniche usate dai baschi in quest'operazione, ma sappiamo che questi stessi balenieri insegnarono le modalità della caccia alla balena ai marinai inglesi e olandesi che introdussero quest'attività sull'isola di Spitsbergen, nell'oceano Artico. Ci sono pervenute relazioni sul procedimento in uso alle stazioni di Spitsbergen. Le balene venivano rimorchiate o calate fin quasi a riva, dove venivano private del grasso, che veniva staccato in strisce e poi fuso. (Nell'acqua poco profonda davanti ai forni sulla Saddle Island ci sono cumuli di detriti di roccia e scivoli formati da pezzi di legno incastrati con cura fra loro. Forse proprio qui i balenieri, armati di grandi coltelli, staccavano il grasso dalle balene.) Le strisce di grasso venivano tagliate in parti più piccole e gettate nei barili, che venivano poi fatti salire con funi su per una rampa di legno sino a grosse caldaie situate dalla parte rivolta verso il mare del muro del forno. A Spitsbergen una fila di caldaie in rame, ciascuna delle quali aveva una capacità variabile fra 180 e 200 litri (grosso modo un barile di pezzi di grasso), veniva posata su cerchi di mattoni, i quali non solo svolgevano una funzione di sostegno ma formavano anche dei focolari. La parte superiore della caldaia si trovava un paio di metri circa al di sopra del suolo. (Dalla parte del muro rivolta verso il mare, a Saddle Island, sono stati riportati in luce i resti di quattro focolari costruiti con roccia e argilla.)

Questa punta di arpone, lunga in origine 40 centimetri, è una delle varie voci dell'equipaggiamento dei balenieri trovate a Red Bay. La lettera *M*, che è probabilmente un marchio del possessore, assomiglia a un monogramma basco del Cinquecento. Arponi, lance e coltelli per staccare il grasso dalle balene uccise erano tra gli oggetti imbarcati in Spagna con destinazione Butus (forse Red Bay).

Un osservatore di questa fase della lavorazione a Spitsbergen, il capitano Thomas Best, riferì sull'uso di collocare dalla parte del muro del forno, rivolta verso l'interno dell'isola, una scialuppa non più idonea alla navigazione, ma dallo scafo ancora impermeabile e di riempirla in parte d'acqua. L'olio veniva poi versato con mestoli dalle caldaie nell'acqua per opera di uomini che stavano su un'impalcatura costruita lungo il muro. L'acqua serviva a raffreddare l'olio bollente e in parte anche a depurarlo; le scorie scendevano sul fondo della scialuppa e l'olio di balena galleggiava.

Dalla parte del muro rivolta verso l'interno a Saddle Island sono stati trovati alcuni resti di un'impalcatura simile, preservati per saturazione. Questa piattaforma di lavoro era stata costruita principalmente con legname locale, ma varie tavole di quercia con buchi dimostravano che i costruttori baschi avevano riutilizzato anche legname già usato per barche o navi. Anziché usare un vecchio scafo come recipiente per l'acqua utilizzata per raffreddare l'olio, qui i balenieri avevano allineato sull'impalcatura mastelli da 90 litri, formati da barili tagliati a metà. Anche le doghe, i pezzi di fondi e i cerchi trovati qui al di sopra del legname di abete e di quercia dell'impalcatura, corrispondevano perfettamente alle parti di barili recuperate dal relitto. Trovammo anche tegole per tetti e parti di un'intelaiatura in legno, a dimostrazione del fatto che gli uomini che versavano l'olio nell'acqua lavoravano al coperto e che anche le caldaie erano protette dalle intemperie.

Il documento che condusse alla ricerca di una nave affondata al largo di Saddle Island registra gli atti di un processo. Esso comprende l'affermazione che, verso la fine del 1565, la nave *San Juan,* dopo avere già imbarcato il suo carico di olio di balena, stava preparandosi a prendere il mare diretta verso la Spagna, quando andò perduta «per mancanza di una gomena». L'espressione suggerisce che una tempesta improvvisa deve avere spezzato gli ormeggi della nave a pieno carico, spingendola verso la riva, dove si infranse. Nel 1978 i subacquei della Marine Excavation Unit avevano visto strutture in legno di una nave emergere dallo strato di fango del fondo a una profondità di 10 metri in prossimità della costa settentrionale di Saddle Island. La prima cosa che essi videro emergendo, dopo aver compiuto un'ispezione sul fondo, fu la trincea di prova scavata dagli archeologi della Memorial University nel sito di uno dei forni sulla riva.

Il fatto che il relitto si trovi così vicino a una stazione costiera basca fece sorgere la speranza che si trattasse effettivamente di una delle navi basche e forse addirittura della *San Juan.* Le ricerche preliminari svolte successivamente dalla Marine Excavation Unit hanno rafforzato questa speranza. La nave è fatta principalmente di quercia, un fatto bastante a escludere che si tratti di una nave posteriore, costruita sul posto. Si aggiunga che il relitto

Muri fatti di pietre locali e di ghiaia furono costruiti con andamento parallelo alla costa a Saddle Island. Le grandi caldaie per la fusione del grasso stavano dalla parte del muro rivolta verso la costa. La trincea dalla parte verso l'interno conteneva frammenti della piattaforma su cui lavoravano gli uomini che toglievano con mestoli l'olio dalle caldaie, trasferendolo poi in acqua per raffreddarlo e depurarlo dalle scorie.

contiene grandi quantità di doghe di barile fatte di quercia e di faggio e pezzi di fondi e cerchi di barile corrispondenti a quelli rinvenuti nell'area imbevuta d'acqua sull'isola. Si tratti o no della *San Juan*, era chiaro che quella nave era associata all'attività della caccia alla balena praticata dai baschi a Red Bay nel Cinquecento.

Nelle stagioni di ricerca successive al 1978 gli archeologi del Parks Canada hanno compiuto un lavoro importante a Red Bay in cooperazione col governo di Terranova e Labrador. I loro sforzi si sono concentrati su due compiti principali: lo scavo del relitto stesso (compreso lo scavo di una trincea dalla poppa della nave in direzione di Saddle Island) e un'ispezione generale sull'area portuale in generale. Nel corso di quest'ispezione - compiuta in collaborazione dai due gruppi di archeologi - è stata rinvenuta una spiaggia a ovest della cittadina che è cosparsa di crani, vertebre, costole e altre ossa di balena boreale. Sulla parte nordoccidentale della Penney's Island sono stati riportati in luce anche i resti di almeno altri due forni.

Prescindendo dalle ossa di balena, le uniche prove di attività dei baschi sulla spiaggia, tanto in superficie quanto in uno scavo di prova, furono alcuni frammenti di tegole. Al largo della costa di Penney's Island, però, i subacquei localizzarono altre due concentrazioni di ossa di balena. Pare che, dopo che le balene erano state private del loro grasso, che veniva trasportato a riva, le loro carcasse venissero lasciate semplicemente affondare.

Al largo dell'estremità occidentale di Penney's Island, i subacquei hanno scoperto i resti di una seconda nave affondata, considerevolmente più piccola di quella di Saddle Island. La sua costruzione e le tecniche di fissaggio suggeriscono che si tratti di una nave di età considerevole, ma solo ulteriori studi potranno dirci se sia o meno contemporanea all'attività di caccia alla balena a Red Bay.

Oltre alle trincee scavate sulla costa, altre tre trincee esplorative furono scavate sott'acqua per determinare i limiti del naufragio. Una di esse si estendeva a nord della prua della nave e le altre due erano dirette verso est e verso ovest, ossia a babordo e a tribordo. Il relitto giace sul suo fianco destro; perciò i resti delle sue strutture e i manufatti furono trovati ampiamente dispersi in tale direzione. Il fondo del porto è formato da fango, roccia e frammenti di fasciame. Scavando la trincea in direzione della costa, il gruppo subacqueo scoprì uno spesso strato di schegge di legno identiche a quelle trovate a migliaia a riva: scarti della lavorazione di legname locale. Poiché le schegge sono affondate anziché galleggiare ed essere quindi trasportate via, si può supporre che esse rappresentino un deposito secondario di rifiuti che si erano già impregnati d'acqua nel sito di costruzione.

Al di sopra dello strato di schegge i subacquei trovarono un secondo strato di materiali organici che in taluni luoghi aveva uno spessore anche di 20 centimetri. Esso era formato da ossa di uccelli, di mammiferi terrestri, di balene e, soprattutto, di merluzzo. Le ossa di pesce erano principalmente crani e frammenti di vertebre; l'assenza di altri resti suggerisce che i baschi aprissero il pesce per salarlo e seccarlo. Questo reperto può spiegare una scoperta che aveva molto incuriosito il gruppo della Memorial University nel corso della sua prima visita a Saddle Island, ossia una serie di piattaforme circolari formate da pietre e ghiaia, del diametro di circa due metri. Può darsi che i marinai vi ammassassero il merluzzo salato dopo l'essiccatura. Nel deposito riportato in luce dalla trincea scavata in direzione dell'isola c'erano, oltre a ossa di balena, anche ossa di merluzzo, a indicazione del fatto che i baschi praticavano simultaneamente alla caccia alla balena, anche la pesca del merluzzo. Fra i manufatti riportati in luce in questa trincea vi sono gli onnipresenti frammenti di tegole, cocci di ceramica, pezzi di funi, un certo numero di scarpe di cuoio ben conservate e i resti parziali di una o due piccole barche baleniere.

L'obiettivo principale dell'indagine subacquea è il relitto stesso. L'interesse particolare di questo reperto risiede non solo nella grande quantità dei resti e nel loro eccellente stato di conservazione, ma anche nel fatto che il legname della nave può fornirci informazioni uniche sulla costruzione delle navi mercantili del Cinquecento. Secondo i documenti studiati dalla Barkham, il *San Juan* naufragò in acque abbastanza poco profonde da consentire all'equipaggio di salvare una parte del carico, una parte dei viveri per il viaggio di ritorno e alcuni fra gli oggetti di proprietà personale. Perciò, se il relitto di Saddle Island è veramente il *San Juan*, si può supporre che le strutture superiori della nave siano rimaste fuor d'acqua, forse per vari anni. In ogni caso l'azione dei ghiacci invernali deve avere accelerato il crollo dello scafo. Depositi di fango e rifiuti, fra cui frammenti della nave stessa, si sono accumulati su entrambi i fianchi del relitto. Questa coltre protettiva, in combinazione con le basse temperature delle acque del Labrador, ha preservato tanto il relitto quanto la parte del suo carico rimasta a bordo.

Il gruppo per la ricerca subacquea ha ora scavato una serie di aree di due metri di lato, soprattutto nella sezione centrale del relitto. Questi scavi hanno riportato in luce l'intera larghezza della nave al di sotto della linea di galleggiamento, davanti e dietro la posizione stimata dell'albero maestro, e anche parti della prua e della poppa. Sin quasi dall'inizio del loro lavoro, i subacquei si sono imbattuti nei resti di barili per il trasporto di olio di balena, rotti, ma virtualmente completi. Alcuni barili erano dispersi a caso, mentre altri si trovavano ancora dov'erano stati stivati, fermati da pietrisco e da corte zeppe di legno per impedire loro di spostarsi. I barili dovevano contenere circa 200 litri d'olio. Al termine della stagione 1980 gli archeologi subacquei avevano riportato in superficie e registrato più di 10 000 frammenti di barile.

La chiglia della nave era lunga circa 15 metri; sulla base di questa dimensione si può stimare la sua capacità di carico a circa 300 tonnellate. La fortissima inclinazione della prua e della poppa, la curva dolce del dritto di prua e il disegno del quadro di poppa sono altrettanti elementi compatibili con l'architettura navale del Cinquecento. Altri elementi riflettono però tecniche di costruzione più antiche. Per esempio l'estremità posteriore della chiglia sale verso l'alto «a ginocchio» ed è fissata al dritto di poppa in un modo simile a quello usato nella costruzione delle navi mercantili medioevali note come cocche. (Recentemente gli archeologi marini hanno scoperto una di queste navi nel porto di Brema.)

Eccezion fatta per i barili, nel relitto sono stati finora trovati pochi altri manufatti. Circa 22 metri a est della chiglia fu scoperta un'ancora, che è attualmente sottoposta a trattamento per eliminare gli effetti dell'immersione prolungata. Sono stati recuperati l'argano, uno dei pochi esemplari noti che risalgano al Cinquecento, e vari bozzelli. Due sezioni di tubi di una pompa di sentina sono stati localizzati, ma non recuperati. Forse il singolo manufatto più interessante è un cannone girevole, su affusto a piedestallo, di ferro battuto, lungo due metri. Questi cannoni a retrocarica, montati sulla battagliola della nave, erano chiamati «assassini» perché spesso arrecavano agli artiglieri un danno non minore che al bersaglio. Un altro enigma è rappresentato dal fatto che il cannone, esaminato dagli archeologi, è risultato carico. Ci si chiede perché.

È in corso il recupero di singoli pezzi di legname della nave, cosicché si potranno cominciare a far copie a ricalco e disegni in scala. Man mano che gli scavi procedono viene registrata l'ubicazione di ciascun elemento strutturale e anche di frammenti della nave naufragata. Le informazioni così ottenute saranno utilizzate per costruire un modello in scala del relitto, nella stessa situazione in cui giace sul fondo del porto. Una volta completati gli scavi, si potrà cominciare a lavorare a un secondo modello in scala, con i vari componenti rimontati per mostrare la forma dello scafo della nave. Questa seconda operazione di montaggio accrescerà la nostra comprensione della progettazione e della tecnologia della costruzione navale che produsse le navi oceaniche che nel Cinquecento portavano nel Nuovo Mondo esploratori, coloni e pescatori europei.

GIOCHI MATEMATICI

di Martin Gardner

Il postulato delle parallele di Euclide e i suoi moderni sviluppi

«Le linee parallele
si incontrano all'infinito!»
andava sempre con forza
ripetendo Euclide.
Finché morì,
e si trovò in quei paraggi:
lì scoprì
che quegli accidenti di cose
divergevano

—PIET HEIN, *Grooks VI*

Gli *Elementi* di Euclide sono noiosi e prolissi, non esplicitano il fatto che due cerchi si possono intersecare, che un cerchio ha un esterno e un interno, che i triangoli possono essere ribaltati e trascurano altri presupposti essenziali per il sistema. Rifacendosi ai canoni moderni, Bertrand Russell poteva definire la quarta proposizione di Euclide «un intreccio di insensatezze» e dichiarare scandaloso che gli *Elementi* fossero ancora utilizzati come libro di testo.

D'altra parte, la geometria di Euclide fu il primo importante tentativo di organizzare la materia in un sistema assiomatico e non sembra giusto incolpare Euclide per non aver individuato prima tutti i miglioramenti poi introdotti da David Hilbert e da altri nel corso della formalizzazione del sistema. L'essersi reso conto del fatto che il suo famoso quinto postulato non era un teorema, ma un assioma che andava accettato senza dimostrazione, è la prova più evidente della sua genialità.

Il modo in cui Euclide formulò il postulato era piuttosto contorto, ma ci si accorse ben presto che poteva avere una forma più semplice: dato un punto su un piano, che non appartenga a una retta data, per esso passa una e una sola retta parallela alla retta data. Non essendo ovvio e intuitivo come gli altri assiomi di Euclide, i matematici tentarono per duemila anni di togliere il postulato riducendolo a teorema dimostrabile sulla base degli altri assiomi. Furono tentate centinaia di prove; alcuni eminenti matematici ritennero anche di essere riusciti nell'intento, ma sempre saltava fuori che in qualche punto della loro dimostrazione era stata posta una premessa che era equivalente al postulato delle parallele o richiedeva il postulato stesso.

Per esempio, è facile dimostrare il postulato delle parallele se si presuppone che la somma degli angoli interni di ogni triangolo equivale a due angoli retti; purtroppo, però, non si può dimostrare questo assunto senza usare il postulato delle parallele. Un'antica falsa dimostrazione, attribuita a Talete di Mileto, si fonda sull'esistenza di un rettangolo, cioè di un quadrilatero con quattro angoli retti; non si può però dimostrare che esistono rettangoli senza usare il postulato delle parallele! Nel XVII secolo, un noto matematico inglese, John Wallis, credette di aver dimostrato il postulato; purtroppo non si era accorto che non si può presupporre che due triangoli siano simili, ma non congruenti, senza il postulato delle parallele. Si potrebbe fare un lungo elenco di altri presupposti, tutti intuitivamente ovvi e tutti equivalenti al postulato delle parallele nel senso che essi non valgono se non vale il postulato.

Agli inizi del XIX secolo, il tentativo di dimostrare il postulato divenne quasi una mania. In Ungheria, Farkas Bolyai dedicò gran parte della sua vita a questo compito, di cui in gioventù aveva spesso discusso con l'amico tedesco Carl Friedrich Gauss. Suo figlio János divenne così ossessionato dal problema da spingere suo padre a scrivergli in una lettera: «Per l'amor di Dio, te ne supplico, smettila. Abbine paura non meno che delle passioni sensuali perché anch'esso può toglierti tutto il tempo che hai e privarti della salute, della pace mentale e della felicità.»

János non la smise e presto si persuase che non solo il postulato era indipendente dagli altri assiomi, ma anche che si sarebbe potuto creare una geometria coerente, assumendo che per il punto passino infinite rette parallele alla retta data. «Ho creato un nuovo universo dal nulla» scrisse con orgoglio a suo padre nel 1823.

Farkas chiese subito a suo figlio di lasciargli pubblicare queste sensazionali asserzioni in appendice a un libro che stava completando. «Se veramente sei riuscito nel tuo intento, è giusto che i tuoi risultati siano resi pubblici senza indugi, e questo per due ragioni: primo, perché le idee passano facilmente da uno all'altro che può anticiparne la pubblicazione; secondo, c'è qualcosa di vero nel fatto che spesso in un'epoca certe cose vengono scoperte in numerosi luoghi diversi, proprio come in primavera le viole appaiono in ogni dove. Anche tutte le lotte scientifiche sono vere e proprie guerre, in cui non so quando arrivi la pace; dovremmo quindi conquistare finché ne siamo capaci, dato che è sempre in vantaggio chi arriva per primo.»

Il breve capolavoro di János apparve in effetti nel libro di suo padre, ma la pubblicazione del libro fu rinviata fino al 1832 e il matematico russo Nikolai Ivanovitch Lobachevski precedette János rivelando in un articolo del 1829 alcuni dettagli della stessa strana geometria (più tardi definita iperbolica da Felix Klein). Quel che è peggio, quando Farkas inviò l'appendice al suo vecchio amico Gauss, il principe dei matematici gli rispose che lodando il lavoro avrebbe lodato soltanto se stesso, visto che egli aveva già da anni sviscerato l'argomento, pur non avendo pubblicato nulla. In altre lettere spiegò che aveva voluto evitare di sollevare le «strida» dei «Beoti», come egli chiamava i suoi colleghi conservatori. (Nell'antica Atene i Beoti erano considerati alquanto stupidi.)

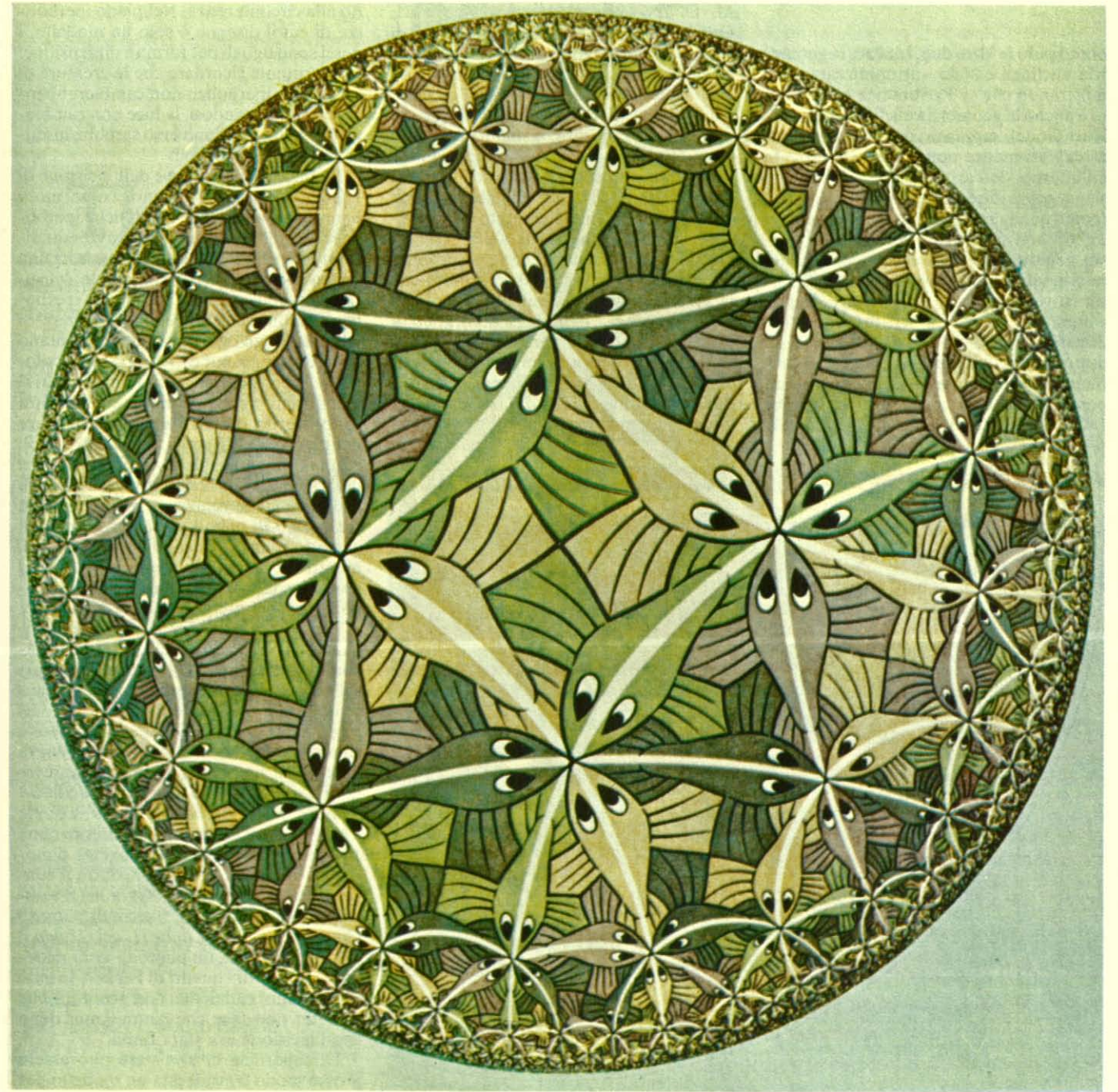
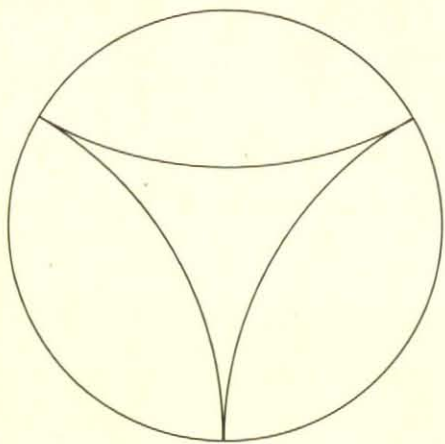
Sconvolto dalla risposta di Gauss, János arrivò a sospettare che suo padre avesse fatto sapere qualcosa della sua meravigliosa scoperta a Gauss. Quando in seguito seppe dell'articolo di Lobachevski, perse qualsiasi interesse per l'argomento e non pubblicò più nulla. «La natura dell'effettiva verità non può essere che unica e uguale in Marcos-Vasarhely come in Kamchatka e sulla Luna» scrisse ormai rassegnato al fatto di aver troppo ritardato la pubblicazione per poter raggiungere l'onore in cui aveva così appassionatamente sperato.

In certa misura, la storia del gesuita italiano Girolamo Saccheri è ancora più triste di quella di Bolyai. Già nel 1733, nel libro *Euclides ab omni naevo vindicatus,* Saccheri aveva costruito senza saperlo entrambi i tipi di geometria non euclidea (il secondo tipo lo vedremo più avanti)! O almeno così sembra. Saccheri rifiutò di credere che le due geometrie fossero coerenti, ma arrivò così vicino ad accettarle che alcuni storici ritengono che egli sostenesse di negarle solo per poter pubblicare il suo libro. «Affermare che un sistema non euclideo era "vero" come quello di Euclide - scrive Eric Temple Bell in un capitolo di *The Magic of Numbers* dedicato a Saccheri - sarebbe stato un temerario invito alla repressione e alla disciplina. Il Copernico della geometria ricorse quindi al sotterfugio: correndo un bel rischio, Saccheri denunciò il proprio lavoro, sperando con questo pietoso tradimento di far sfuggire ai censori il suo passato eretico.»

Vorrei aggiungere due aneddoti sui Bolyai. János era un ufficiale di cavalleria (la matematica era sempre stata rigorosamente uno svago) noto per la sua abilità nella scherma, la bravura come violinista e il temperamento ardente. Si dice che avesse sfidato a duello tredici ufficiali, alla condizione che, dopo ogni vittoria, gli fosse consentito suonare il violino per lo sconfitto. Bolyai padre pare sia stato sepolto su sua richiesta sotto un albero di mele, senza nessun monumento, per commemorare le tre più famose mele della storia: la mela di Eva, la mela d'oro data da Paride a Venere come premio in una gara di bellezza e la mela ispiratrice di Isaac Newton.

Prima della fine del XIX secolo, divenne chiaro che il postulato delle parallele non solo era indipendente dagli altri, ma poteva anche essere modificato in due modi opposti. Se veniva sostituito (come proposto da Gauss, Bolyai e Lobachevski) dall'assunzione che un numero infinito di rette «ultraparallele» passino per il punto, il risultato era una nuova geometria altrettanto elegante e «vera» di quella di Euclide. Tutti gli altri postulati di Euclide rimangono validi; una linea «retta» è ancora una geodetica, ossia la linea più breve. In questo spazio iperbolico tutti i triangoli hanno una somma degli angoli inferiore a 180 gradi e la somma diminuisce con l'allargarsi del triangolo; tutti i poligoni simili sono congruenti; la circonferenza di qualsiasi cerchio è maggiore di π volte il diametro; la misura di curvatura del piano iperbolico è negativa (mentre la curvatura del piano euclideo è zero) e uguale in qualsiasi punto; come la geometria euclidea, la geometria iperbolica è generalizzabile a un spazio a tre e più dimensioni.

Il secondo tipo di geometria non euclidea, definita da Klein «ellittica», fu sviluppato più tardi e simultaneamente dal matematico tedesco Georg Friedrich Bernhard Riemann e dal matematico svizzero Ludwig Schläfli. In esso il postulato delle parallele viene sostituito dall'assunto che per il punto non passi *nessuna* linea parallela alla linea data. In questa geometria la somma degli angoli di un triangolo è sempre maggiore di 180 gradi e la circonferenza di un cerchio è sempre minore di π volte il diametro; tutte le geodetiche sono finite e chiuse; ogni coppia di geodetiche si interseca.

Per provare la coerenza delle due nuove geometrie furono trovati per entrambe dei modelli euclidei che dimostravano che se la geometria euclidea era coerente lo

Limite III del cerchio, *di M. C. Escher, è una tassellatura basata su un modello euclideo del piano iperbolico*

Il più grande triangolo sul piano iperbolico

erano pure le altre due. Inoltre, la geometria euclidea è stata «aritmetizzata», dimostrando che se l'aritmetica è coerente lo è anche la geometria euclidea. Grazie a Kurt Gödel, sappiamo ora che la coerenza dell'aritmetica non può essere provata all'interno dell'aritmetica e sebbene vi siano prove di coerenza per l'aritmetica (come quella famosa data da Gentzen nel 1936), non si è trovata una prova di questo genere che possa essere considerata interamente costruttiva da un intuizionista (si veda l'articolo *Constructive Mathematics*, di Allan Calder in «Scientific American», ottobre 1979). Dio esiste, ha detto una volta qualcuno, perché la matematica è coerente; e il diavolo esiste perché non siamo capaci di dimostrarlo.

Come ha detto Paul C. Rosenbloom, le varie metadimostrazioni della coerenza della matematica non hanno forse eliminato il diavolo, ma hanno ridotto quasi a zero le dimensioni dell'inferno. È certo che nessun matematico oggi si aspetta che l'aritmetica (quindi anche le geometrie euclidea e non euclidea) producano una contraddizione. Curiosamente, Lewis Carroll fu uno degli ultimi matematici a dubitare della geometria non euclidea. «È strano - ha scritto lo studioso di geometria H. S. Coxeter - che proprio lui, la cui Alice nel paese delle meraviglie era in grado di modificare la propria dimensione mangiando semplicemente un dolce, fosse incapace di accettare la possibilità che l'area di un triangolo possa rimanere finita quando i lati tendono all'infinito.»

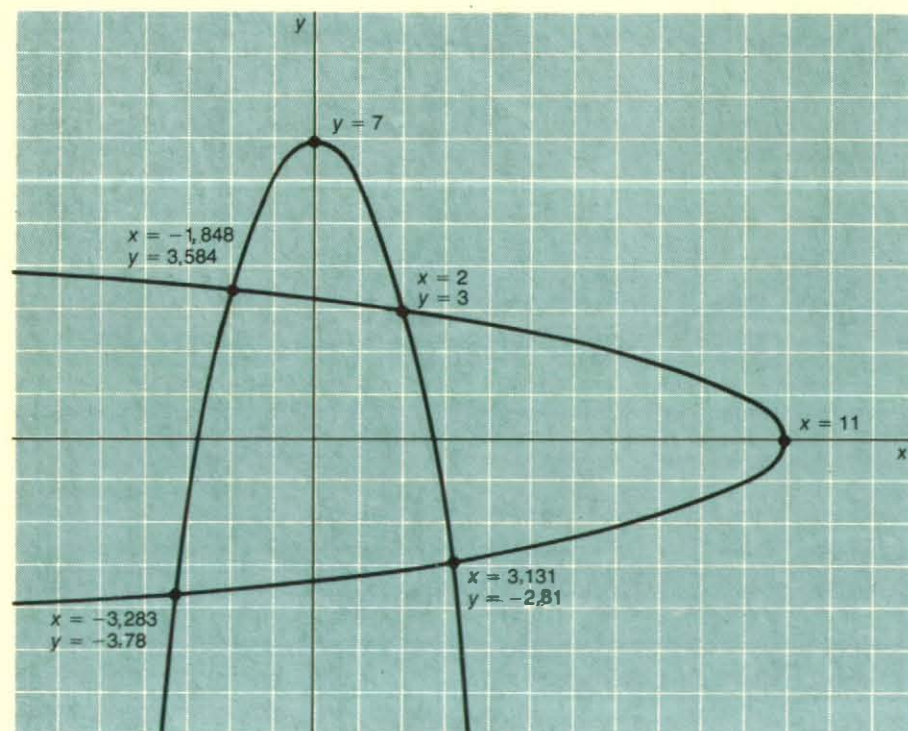
Per capire ciò a cui si riferiva Coxeter si può esaminare il *Limite III del cerchio* di M. C. Escher riprodotto a pagina 111. Questa xilografia (uno dei pochi lavori di Escher in cui vengono utilizzati vari colori) è una tassellatura basata su un modello euclideo, costruito da Henry Poincaré, del piano iperbolico. Nell'ingegnoso modello di Poincaré, a ogni punto del piano euclideo corrisponde un punto interno (ma non sopra) alla circonferenza del cerchio. Al di là del cerchio c'è, come dice Escher, il «nulla assoluto».

Si immagini che degli abitanti piatti vivano su questo modello. Se si spostano dal centro verso l'esterno, ci sembra che la loro dimensione divenga sempre più piccola, anche se essi non si accorgono di alcun cambiamento perché anche i loro strumenti di misura divengono progressivamente più piccoli. Al confine la loro dimensione diventa zero, ma non possono mai raggiungere il confine stesso. Se procedono verso di esso a velocità costante, la loro velocità (per noi) si riduce progressivamente, anche se a loro sembra costante. La luce iperbolica segue delle geodetiche, ma poiché la sua velocità è proporzionale alla sua distanza dal confine essa segue dei percorsi che ci appaiono come degli archi circolari che incontrano il confine secondo angoli retti.

In questo mondo iperbolico un triangolo ha un'area massima finita, come si vede nella figura in alto di questa pagina, anche se i suoi tre lati «diritti» vanno all'infinito in lunghezza iperbolica e i suoi tre angoli sono pari a zero. Non si deve immaginare che il mosaico di Escher sia disposto su una sfera: è un cerchio che racchiude un'infinità di pesci - Coxeter lo chiama una «pesca miracolosa» - che diventano sempre più piccoli man mano si avvicinano alla circonferenza. Nel piano iperbolico, di cui il disegno è solo un modello, i pesci sono uguali per forma e dimensione. È importante ricordare che le creature di un mondo iperbolico non cambierebbero di forma spostandosi, la luce non cambierebbe velocità e l'universo sarebbe infinito in tutte le direzioni.

Le linee bianche curve dell'incisione di Escher non rappresentano, come molti hanno supposto, delle geodetiche iperboliche. Le linee sono dette curve equidistanti di ipercicli e ciascuna di esse ha una distanza perpendicolare costante (misurata iperbolicamente) dalla retta iperbolica che unisce le estremità dell'arco. Si noti che lungo ogni curva bianca nuotano coda contro testa pesci dello stesso colore. Se si considerano tutti i punti in cui si incontrano quattro pinne, questi punti sono i vertici di una tassellatura regolare del piano iperbolico da parte di triangoli equilateri con angoli di 45 gradi. I centri dei triangoli sono i punti in cui tre pinne di sinistra si incontrano e tre bocche toccano tre code. Gli angoli di 45 gradi fanno sì che otto triangoli circondino ogni vertice, mentre in una tassellatura euclidea da parte di triangoli equilateri solo sei triangoli possono circondare ogni vertice.

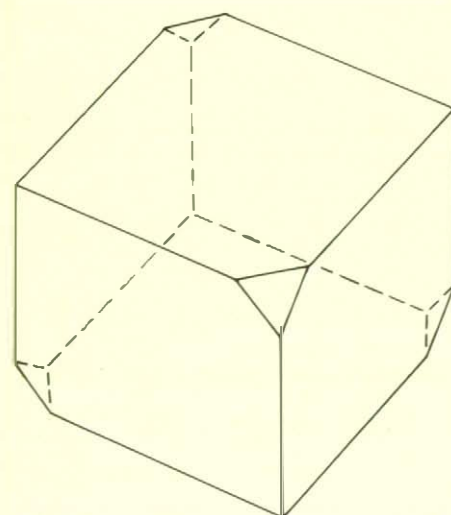
Escher e Coxeter erano in corrispondenza dal 1954, anno in cui si erano incontrati, e l'interesse di Escher per le tassellature del piano iperbolico era aumentato dopo aver visto le figure che accompagnavano un articolo del 1957 sulla simmetria dei cristalli che Coxeter aveva scritto e gli aveva inviato. In un incantevole articolo intitolato *The Non-Euclidean Symmetry of Escher's Picture «Circle Limit III»* (nella rivista «Leonardo», vol. 12, pagine 19-25, 1979) Coxeter dimostra che ogni arco bianco incontra il confine a un angolo di circa 80 gradi. (Il valore preciso è $2^{7/4} + 2^{5/4}$ secondi d'arco.) Coxeter considera *Limite III del cerchio* il più sofisticato da un punto di vista matematico tra tutti i quadri di Escher. In esso era perfino anticipata una scoperta che Coxeter non fece che cinque anni dopo che l'incisione era stata finita.

La superficie di una sfera può essere grosso modo considerata un modello per la geometria ellittica. Qui le rette euclidee divengono cerchi massimi. È chiaro che

Risposta al problema 7-11

non ve ne possono essere due paralleli ed è facile capire che i triangoli formati da archi dei cerchi massimi devono avere angoli con somma superiore a due angoli retti. Analogamente, il piano iperbolico può trovare un modello nella superficie a forma di sella di una pseudosfera generata dalla rotazione di una trattrice lungo il suo asintoto.

È scorretto accusare di insensatezza quei matematici che, prima che si fosse stabilita l'indipendenza del postulato delle parallele, ritenevano erroneamente di aver dimostrato il postulato stesso. Lo stesso non si può dire per quei dilettanti dei decenni successivi che non riuscivano a capire le prove di indipendenza del postulato o erano troppo egocentrici per sforzarsi di capirle. Augustus De Morgan, nel suo classico compendio di eccentricità matematiche, *A Budget of Paradoxes*, ci fa conoscere il più infaticabile dimostratore del postulato delle parallele del XIX secolo, il generale Perronet Thompson. Thompson continuava a sfornare revisioni delle sue molte prove (una era basata sulla spirale equiangolare) e De Morgan non riuscì mai nell'intento di dissuaderlo dai suoi futili sforzi. Thompson voleva anche rimpiazzare la scala temperata del pianoforte con un'ottava di 40 note.

Il più ridicolo dei dimostratori americani del postulato delle parallele fu il Molto Reverendo Jeremiah Joseph Callahan, rettore della Duquesne University di Pittsburgh. Nel 1931, quando Padre Callahan annunciò di aver trisecato l'angolo, il «Time» diede un breve resoconto della storia e pubblicò una sua fotografia. L'anno seguente, Callahan pubblicò il suo libro più importante, *Euclid or Einstein: A Proof of the Parallel Theory and a Critique of Metageometry* (Devon-Adair, 1932), un trattato di 310 pagine in cui parlava di Einstein in questi termini: è «confuso», «non ha una mente logica», si trova in una «nebbia mentale», è un «pensatore trascurato». «A volte viene da ridere - scriveva Callahan - e a volte ci si sente irritati... Ma non c'è da aspettarsi che Einstein ragioni.»

Nuovo rompicapo sulla dissezione di una scacchiera

Cubo troncato che obbliga all'uso di sei colori

Ciò che Callahan trovava così irritante era l'adozione da parte di Einstein di una geometria non euclidea generalizzata, formulata da Riemann, in cui la curvatura dello spazio fisico varia da punto a punto a seconda dell'influenza della materia. Una delle grandi rivoluzioni della teoria della relatività era la scoperta che si poteva ottenere un'enorme e generale semplificazione della fisica assumendo che lo spazio fisico abbia questo tipo di struttura non euclidea.

È ora un luogo comune (quanto sarebbe rimasto colpito e deliziato Kant da questa nozione!) riconoscere che tutti i sistemi geometrici sono ugualmente «veri» in astratto, ma che la struttura dello spazio fisico deve essere determinata empiricamente. Lo stesso Gauss pensava di triangolare tre vette montuose per vedere se i loro angoli avevano per somma due angoli retti; si dice anzi che abbia davvero effettuato questa prova, con risultati non conclusivi. Sebbene si possa dimostrare sperimentalmente che lo spazio fisico è non euclideo, è curioso che non ci sia modo di dimostrare che esso è euclideo! La curvatura zero è un caso limite, a metà tra la curvatura ellittica e quella iperbolica. Dato che tutte le misure sono soggette a errore, la deviazione dallo zero potrebbe essere troppo leggera per essere individuata.

Poincaré era dell'opinione che se gli esperimenti ottici sembravano dimostrare che lo spazio fisico non è euclideo, sarebbe stato meglio mantenere la più semplice geometria non euclidea e supporre che i raggi della luce non seguissero delle geodetiche. Molti matematici e fisici erano d'accordo con Poincaré, compreso Russell, finché la teoria della relatività non modificò le loro opinioni. Alfred North Whitehead fu tra i pochi a non cambiare mai opinione; scrisse perfino un libro sulla relatività, ora dimenticato, in cui si schierava per la conservazione di un universo euclideo (o almeno un universo a curvatura costante) e la modificazione secondo le necessità delle leggi fisiche. A questo proposito si veda Robert M. Palter, *Whitehead's Philosophy of Science*, University of Chicago Press, 1960.

I fisici non sono più turbati dalla nozione che lo spazio fisico abbia una struttura non euclidea generalizzata. Callahan non era solo turbato; era anche convinto che tutte le geometrie non euclidee siano autocontradditorie. Einstein, povero diavolo, non sapeva quanto fosse facile dimostrare il postulato delle parallele. Se vi incuriosisce sapere qualcosa della dimostrazione di Callahan e del suo elementare errore, si veda l'articolo di D. R. Ward in «The Mathematical Gazette» (Vol. 17, pagine 101-104, maggio 1933).

Come i loro cugini che trisecano l'angolo, quadrano il cerchio e trovano semplici dimostrazioni dell'ultimo teorema di Fermat, i dimostratori del postulato delle parallele sono gente decisa. Un esempio recente è William L. Fischer, di Monaco, che nel 1959 ha pubblicato una *Critica della geometria non euclidea* in 100 pagine. Ian Stewart illustra i suoi errori sulla rivista inglese «Manifold» (n° 12, pagine 14-21, estate 1972) e cita una lettera in cui Fischer accusa l'*establishment* matematico di tener nascosto il suo grande lavoro e i giornali ortodossi di rifiutarsi di recensirlo: «La biblioteca dell'Università di Cambridge ha perfino rifiutato di mettere in catalogo il mio libro... Ho dovuto scrivere al vice cancelliere per superare questo boicottaggio.»

Naturalmente non ci sono criteri precisi per distinguere i matematici incompetenti da quelli bravi, ma se è per questo non ci sono nemmeno criteri precisi per distinguere il giorno dalla notte, la vita dalla non vita e per capire dove finisce l'oceano e inizia la costa. Se non avessimo parole per indicare parti dei vari continui non potremmo pensare né parlare. Se tu, caro lettore, conosci un modo per dimostrare il postulato delle parallele, non me ne venire a parlare!

Il primo problema del mio articolo del dicembre scorso richiedeva di fornire con l'uso di una parabola una breve dimostrazione del numero di soluzioni per la coppia di equazioni: $x^2 + y = 7$ e $x + y^2 = 11$. Le equazioni sono rappresentate graficamente da due parabole incrociate, come si vede nella figura in basso di pagina 112. Dato che le parabole si intersecano esattamente in quattro punti, ci sono esattamente quattro soluzioni, di cui solo una ($x = 2, y = 3$) è un intero. Anche se la rappresentazione grafica è imprecisa, si può dimostrare che le altre tre soluzioni non sono numeri interi, verificando i numeri indicati dagli incroci del reticolo più vicini alle intersezioni. In effetti i numeri, tutti irrazionali, hanno i valori approssimati dati nella figura.

Nel secondo problema, posto da Ronald L. Graham, si chiedeva come fosse possibile disporre sul piano un numero infinito di dischi identici con diametro inferiore a 1, poniamo 1/10, in modo che non vi siano due punti con distanza pari a un intero. Una soluzione consiste nel disporli con i centri su una parabola della forma $y = x^2$, con i centri nei punti (1,1), (3,9), (9,81),..., $(3^k,3^{2k})$,... Mi manca lo spazio per dare la dimostrazione di Graham, non pubblicata, e sebbene riluttante devo quindi lasciarla come esercizio per i lettori interessati all'argomento.

Uno dei brevi problemi della mia rubrica di giugno 1981 chiedeva di tagliare una

Mappa a venti regioni che obbliga all'uso di sette colori

coperta patchwork di dimensioni 9 per 12, con un buco 1 per 8, in due parti che insieme formino una coperta senza buchi di dimensione 10 per 10. Eric Scott mi ha scritto per sottopormi il seguente problema: se la coperta rovinata è colorata come una scacchiera, formando il quadrato la colorazione risulta distrutta; come si può tagliare la coperta nel minor numero possibile di pezzi in modo da formare un quadrato che conservi la colorazione?

Il problema è risultato interessante. Ci sono centinaia di soluzioni con quattro pezzi, ma infine ne ho trovata una (che credo fondamentalmente unica) con solo tre pezzi. Tale soluzione può essere generalizzata a tutti i quadrati con lati pari e (con piccole variazioni) a tutti i quadrati di ordine dispari. Se n è il lato del quadrato ed è maggiore di 2, supponiamo che il rettangolo sia $(n - 1) \times (n + 2)$, con un buco di $1 \times (n - 2)$ parallelo al lato lungo del rettangolo e il più centrale possibile.

Propongo ora il nuovo problema con n uguale a 8. Riuscite a tagliare lungo le linee del reticolo il rettangolo mutilato, che si vede nella figura in alto della pagina a fronte, in tre pezzi che formino una scacchiera perfetta? È consentito ruotare i pezzi e rovesciarli. Nel mio prossimo articolo darò la soluzione.

Il problema di Steven J. Brams relativo al gioco sulla colorazione delle mappe, anch'esso nella mia rubrica di giugno, ha portato ad alcuni sorprendenti risultati Robert High, un matematico che lavora per la Informatics, Inc., di New York. Chiamiamo Min il primo giocatore (quello che cerca di minimizzare i colori) e Max il secondo giocatore (quello che cerca di massimizzare i colori). Avevo dato una mappa di sei regioni su cui Max poteva obbligare a usare cinque colori. High ha scoperto che una proiezione del cubo dà una più semplice mappa di sei regioni per obbligare all'uso di cinque colori: Max deve mettere a ogni turno un nuovo colore sulla faccia opposta a quella su cui Min ha effettuato l'ultima giocata.

La mappa per costringere a usare sei colori, che ho dato in agosto, era una proiezione dello scheletro di un dodecaedro. High ha scoperto che se si rimpiazzano quattro angoli di un cubo con facce triangolari, come si vede nella figura in basso della pagina a fronte, una proiezione dà una mappa di sole 10 regioni su cui Max può obbligare a usare sei colori. La strategia è più complicata di quella per la mappa a 12 regioni ricavata dal dodecaedro e quindi non la fornirò.

La maggiore sorpresa è data dalla scoperta di High di una mappa a 20 regioni su cui Max può obbligare a usare sette colori! Si immagini di sostituire ciascun angolo di un tetraedro con una faccia triangolare, e poi di sostituire ciascuno dei 12 nuovi angoli con una faccia triangolare. Nella figura in alto si vede una proiezione piana dello scheletro del poliedro risultante. Qui la strategia è più complessa di quella per la mappa a 10 regioni, ma High mi ha inviato una rappresentazione del gioco in forma ad albero che fornisce la dimostrazione. High suppone che nessuna mappa piana consenta a Max di obbligare all'uso di più di sette colori, ma la cosa rimane indimostrata. È anche possibile che non ci sia un limite superiore minimo.